



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 113470

TO: Georgia L Helmer
Location: REM/2C15/2C18
Art Unit: 1638
Thursday, February 05, 2004

Case Serial Number: 09643755

From: Mary Jane Ruhl
Location: Biotech-Chem Library
Remsen 1-B55
Phone: 571-272-2524

maryjane.ruhl@uspto.gov

Search Notes

Examiner Helmer,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl
Technical Information Specialist
STIC
CM-1, Rm. 6-A-06
605-1155



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2004, 01:29:45 ; Search time 4588 Seconds

(without alignments)
10459.240 Million cell updates/sec

Title: US-09-643-755B-1

Perfect score: 1173

Sequence: 1 atgaacttcctaagctctt.....ggctagctaagaactctga 1173

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_scs: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_scs: *
28: em_un: *
29: em_vl: *
30: em_hcg_hum: *
31: em_hcg_inv: *
32: em_hcg_other: *
33: em_hcg_mus: *
34: em_hcg_pln: *
35: em_hcg_rtd: *
36: em_hcg_mam: *
37: em_hcg_vrt: *
38: em_hy: *
39: em_hgo_hum: *
40: em_hgo_mus: *
41: em_hgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1173	100.0	1173	6	AX088019 Sequence
2	1173	100.0	3957	6	AX088021 Sequence
3	900.4	76.8	1305	4	BOVCHYMOR
4	897.2	76.5	1240	6	AR002347
5	897.2	76.5	1460	6	E00042 DNA coding
6	896.2	76.4	1291	6	A15836 chymosin ge
7	895.6	76.4	1175	6	I04058 Sequence 5
8	894.8	76.3	2733	6	AR073077 Sequence
9	894	76.2	1275	6	BOVCHYMOA
10	894	76.2	1290	6	E00075
11	894	76.2	1291	6	A15633
12	893	76.1	1289	6	E00144
13	893	76.1	2726	6	E08097
14	892.4	76.1	1311	6	E00108
15	892.4	76.1	1460	6	E00295
16	890.8	75.9	1269	4	BOVCHYMOC
17	886.8	75.6	1098	6	E00079
18	886.8	75.6	1098	6	E02341
19	872.8	74.4	2982	6	I08098
20	870.8	74.2	1101	4	AF177290
21	839.6	71.6	1292	4	OAPCHY
22	776	66.2	1115	12	SYNPROCA
23	753.6	64.2	1094	4	BP19786
24	749	63.9	1172	4	CDR131677
25	732.4	62.4	1230	4	SSU14406
26	718	61.2	1262	9	AB038386
27	702	59.8	1095	4	AF421161
28	681.6	58.1	1080	4	AF421162
29	627.2	53.5	969	6	A02000 Artificial
30	627.2	53.5	969	6	A06446 Artificial
31	619.6	52.8	1146	12	SYNPROCA
32	605	51.6	1259	10	RN0251688
33	569	48.5	957	4	AF421164
34	534	45.5	1140	6	AX319837
35	490.8	41.8	981	4	AF421163
36	382.4	32.5	637	6	E05472
37	381.8	32.5	1403	4	AB047246
38	378.6	32.3	1161	4	AB047245
39	373.8	31.9	858	4	AF421165
40	373.8	31.9	1351	9	AB038384
41	370.8	31.6	1381	4	AB047244
42	363.8	31.0	1362	4	AB047243
43	363.8	31.0	1363	4	PICPERA
44	363.4	31.0	1221	4	CDR131678
45	362.2	30.9	1276	5	CHKPSN

ALIGNMENTS

RESULT 1
AX088019 1173 bp DNA linear PAT 17-MAR-2001
LOCUS Sequence 1 from Patent WO0114571.
ACCESSION AX088019
VERSION AX088019.1 GI:13396947
KEYWORDS
SOURCE
ORGANISM
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS
TITLE
van Rooijen, G., Keon, R. G., Boothe, J. and Shen, Y.
Commercial production of chymosin in plants

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0114571-A 1 01-MAR-2001;
SemBioSys Genetics Inc. (CA)
FEATURES Location/Qualifiers
source 1..1173
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
1..1173
CDS
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC34680.1"
/db_xref="GI:13396948"
/translation="MNFLKSPFPYAFLCRGYFVAVTAAETTRIPLYGKSLRKALK
EHLLEDFLOKQOYGISSKTSGGEVAVPLTNVLDQSFGKTYLGPPEFTVLDY
GSSDFWVPSIYCKSNACKMHPDPKRSFQMLGKPLSHVGTSMQGLIGDYTV
SNIVDIQOFTVGLSTOBPGDVFYAEFDGILGMAYPSLASEYSIPVEDNMNRHLVAD
LFSVMDRNGOESMLTLGALIDPSYVGSILHWPVTVQOYQFTVDSVTLSGVVACEG
GCCAIIIDTSTSLKVPSSDIINIQAIGATONOGEPDIDCDNLSTMPVVPFINCKM
YPLTBSAYTSQDGFTSGFOSBNHSQKMLIDGVFIREYISVFDRANMLVGLAKAI"
BASE COUNT 299 a 308 c 262 g 304 t
ORIGIN
Query Match 100.0%; Score 1173; DB 6; Length 1173;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAACCTCTTAAGTCTTTCCCTTTCTACGCTTTCCTTTGTTCCGTCGAATCTTCGTT 60
DB 1 ATGAACCTCTTAAGTCTTTCCCTTTCTACGCTTTCCTTTGTTCCGTCGAATCTTCGTT 60
QY 61 GCTGTACTACAGCTGCTGATGATCACCCGCAATTCCTTCTCAAGAAAGTAACTCTCCGT 120
DB 61 GCTGTACTACAGCTGCTGATGATCACCCGCAATTCCTTCTCAAGAAAGTAACTCTCCGT 120
QY 121 AAGGCGCTGAAGAACATGACTTCTTGAAGACTTCTTGGCAAAACAAGATAGCATC 180
DB 121 AAGGCGCTGAAGAACATGACTTCTTGAAGACTTCTTGGCAAAACAAGATAGCATC 180
QY 181 AGCAGCAAGTACTCCGGCTTCGGTGAAGTGTACGCTGCACCTTCAACCACTACCTTGAT 240
DB 181 AGCAGCAAGTACTCCGGCTTCGGTGAAGTGTACGCTGCACCTTCAACCACTACCTTGAT 240
QY 241 AGTCATATCTTTGGGAAGATCTACCTCGGAACCCCGCTCAAGAGTTACCGTTCTTTT 300
DB 241 AGTCATATCTTTGGGAAGATCTACCTCGGAACCCCGCTCAAGAGTTACCGTTCTTTT 300
QY 301 GATACGTCTCTCTGACTCTGAGTTCCTCTATCTACTGCAAGAGCAATGCTGCAAG 360
DB 301 GATACGTCTCTCTGACTCTGAGTTCCTCTATCTACTGCAAGAGCAATGCTGCAAG 360
QY 361 AACCCAAAGATTCGATCCGAGAAAGTGTCCAACCTTCCGAACCTTAGCCAAACCTTTG 420
DB 361 AACCCAAAGATTCGATCCGAGAAAGTGTCCAACCTTCCGAACCTTAGCCAAACCTTTG 420
QY 421 TCTATACACTACGGGTACAGGTAGCAAGAACTTGAAGCTATGATACCGTCACTGTC 480
DB 421 TCTATACACTACGGGTACAGGTAGCAAGAACTTGAAGCTATGATACCGTCACTGTC 480
QY 481 TCCAAATGTTGACATTTCAACAGACAGTAGGACTTGAACCCCAAGAACAGAGTAGTGC 540
DB 481 TCCAAATGTTGACATTTCAACAGACAGTAGGACTTGAACCCCAAGAACAGAGTAGTGC 540
QY 541 TTCACCTATGACAAATTCGATGCGATCCTTGTATGACATACCATCGCTCGCTGAGAG 600
DB 541 TTCACCTATGACAAATTCGATGCGATCCTTGTATGACATACCATCGCTCGCTGAGAG 600
QY 601 TACTCGATACCTGTGTTGAACAATATGAAACCGAACAAGTAAAGTCAAGACTTGTTC 660
DB 601 TACTCGATACCTGTGTTGAACAATATGAAACCGAACAAGTAAAGTCAAGACTTGTTC 660
QY 661 TCGGTTTACATGAGCAGGAATGGCCAGAGAGCATGCTCAAGCTTGAGCTATTGATCCA 720
DB 661 TCGGTTTACATGAGCAGGAATGGCCAGAGAGCATGCTCAAGCTTGAGCTATTGATCCA 720

QY 721 TCTTACTACACAGGATCTCTTCACTGGGTTTCCAGTCACTGTGAGAGTACTGGCAATTC 780
DB 721 TCTTACTACACAGGATCTCTTCACTGGGTTTCCAGTCACTGTGAGAGTACTGGCAATTC 780
QY 781 ACTGTGAGACAGTGTCAACATACGCGGTGTGTTGTTGCAATGGAAGTGGATGTCAAGCT 840
DB 781 ACTGTGAGACAGTGTCAACATACGCGGTGTGTTGTTGCAATGGAAGTGGATGTCAAGCT 840
QY 841 ATCTTGATACCGGTAAGTGTCAAGCTGTGTCGACCTTAGACGACATTCATCAATTCAG 900
DB 841 ATCTTGATACCGGTAAGTGTGTGAGATCAAGGCAAGATGTACCACTGACCCCTTC 1020
QY 901 CAAGCTATTGAGACCAACAAGAACAGTACGTTGATGACATTAATTGCCAATCTT 960
DB 901 CAAGCTATTGAGACCAACAAGAACAGTACGTTGATGACATTAATTGCCAATCTT 960
QY 961 AGCTATACGCTTCACTGCTTCTTGGATCAACCGCAAGATGTACCACTGACCCCTTC 1020
DB 961 AGCTATACGCTTCACTGCTTCTTGGATCAACCGCAAGATGTACCACTGACCCCTTC 1020
QY 1021 GCCTATACGACGACGATCAAGGTTCTGCACAGTGTATCCAGATGAGAACCATTC 1080
DB 1021 GCCTATACGACGACGATCAAGGTTCTGCACAGTGTATCCAGATGAGAACCATTC 1080
QY 1081 CAGAAATGATCTTGGGAAGTGTTCATTCTGATCTACAGCGTCTTTGACAGGACC 1140
DB 1081 CAGAAATGATCTTGGGAAGTGTTCATTCTGATCTACAGCGTCTTTGACAGGACC 1140
QY 1141 AACACCTGTTGGGCTAGCTAAAGCAATCTGA 1173
DB 1141 AACACCTGTTGGGCTAGCTAAAGCAATCTGA 1173
RESULT 2
AX088021
LOCUS AX088021 3957 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 3 from Patent WO0114571.
ACCESSION AX088021
VERSION AX088021.1 GI:13396949
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS van Rooijen, G., Keon, R. G., Boothe, J. and Shen, Y.
TITLE Commercial production of chymosin in plants
JOURNAL Patent: WO 0114571-A 3 01-MAR-2001;
SemBioSys Genetics Inc. (CA)
FEATURES Location/Qualifiers
source 1..3957
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
1554..2726
/note="Figure 2"
/codon_start=1
/translation="MNFLKSPFPYAFLCRGYFVAVTAAETTRIPLYGKSLRKALK
EHLLEDFLOKQOYGISSKTSGGEVAVPLTNVLDQSFGKTYLGPPEFTVLDY
GSSDFWVPSIYCKSNACKMHPDPKRSFQMLGKPLSHVGTSMQGLIGDYTV
SNIVDIQOFTVGLSTOBPGDVFYAEFDGILGMAYPSLASEYSIPVEDNMNRHLVAD
LFSVMDRNGOESMLTLGALIDPSYVGSILHWPVTVQOYQFTVDSVTLSGVVACEG
GCCAIIIDTSTSLKVPSSDIINIQAIGATONOGEPDIDCDNLSTMPVVPFINCKM
YPLTBSAYTSQDGFTSGFOSBNHSQKMLIDGVFIREYISVFDRANMLVGLAKAI"
BASE COUNT 1263 a 790 c 609 g 1295 t
ORIGIN
Query Match 100.0%; Score 1173; DB 6; Length 3957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACCTCTTAAGTCTTCCCTTCTACGGCTTCTTTGGTGGTCAATACTTCTT 60
 Db 1554 ATGAACCTCTTAAGTCTTCCCTTCTACGGCTTCTTTGGTGGTCAATACTTCTT 1613
 QY 61 GCTGTACTCAACGCTGTGAGATCACCGCATTCCTCTTCAAGAGTACTCTCCGT 120
 Db 1614 GCTGTACTCAACGCTGTGAGATCACCGCATTCCTCTTCAAGAGTACTCTCCGT 1673
 QY 121 AAGGGCGTGAAGAAATGAGCTTCTAGAGCTTCTTGGAGAAACAAGATGAGCATC 180
 Db 1674 AAGGGCGTGAAGAAATGAGCTTCTAGAGCTTCTTGGAGAAACAAGATGAGCATC 1733
 QY 181 AGCAGCAAGTACTCCGGCTCGGTGAAGTGTGCTAGCGCCTTACCAACTCTTAT 240
 Db 1734 AGCAGCAAGTACTCCGGCTCGGTGAAGTGTGCTAGCGCCTTACCAACTCTTAT 1793
 QY 241 AGTCAMTACTTGGGAGAGTCTACCTCGGAACCCCGCTCAAGAGTTCAACGCTTCTT 300
 Db 1794 AGTCAMTACTTGGGAGAGTCTACCTCGGAACCCCGCTCAAGAGTTCAACGCTTCTT 1853
 QY 301 GATATGCTTCTCTGACTTCTGGGTTCCCTTACTTACTGCAAGAGATGCTTGCAG 360
 Db 1854 GATATGCTTCTCTGACTTCTGGGTTCCCTTACTTACTGCAAGAGATGCTTGCAG 1913
 QY 361 AACCAACCAAGTGTGATCCGGAAGTGTGACCTTCCAGACTTATGAGCAACCCCTT 420
 Db 1914 AACCAACCAAGTGTGATCCGGAAGTGTGACCTTCCAGACTTATGAGCAACCCCTT 1973
 QY 421 TCTATACACTACGATACAGTATGATGACAGGATCTTATGAGTATGATCCTGCTGTC 480
 Db 1974 TCTATACACTACGATACAGTATGATGACAGGATCTTATGAGTATGATCCTGCTGTC 2033
 QY 481 TCCAAATTTGTGACATTCACAGACAGTAGSACTTATGACCCCAAGAACAGGTATGTC 540
 Db 2034 TCCAAATTTGTGACATTCACAGACAGTAGSACTTATGACCCCAAGAACAGGTATGTC 2093
 QY 541 TTCACCTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
 Db 2094 TTCACCTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2153
 QY 601 TACTGATACCTGTGTGTGACAAATGATGAAACCAACCTTATGATGATGATGATGATG 660
 Db 2154 TACTGATACCTGTGTGTGACAAATGATGAAACCAACCTTATGATGATGATGATGATG 2213
 QY 661 TCGGTTTACATGACAGAGATGCGCAGAGAGATGCTTCAAGCTTATGATGATGATG 720
 Db 2214 TCGGTTTACATGACAGAGATGCGCAGAGAGATGCTTCAAGCTTATGATGATGATG 2273
 QY 721 TCCCTACTACAGAGATCTTCTTCACTGGGTTCCAGTCACTGTGACAGATGATGATGATG 780
 Db 2274 TCCCTACTACAGAGATCTTCTTCACTGGGTTCCAGTCACTGTGACAGATGATGATGATG 2333
 QY 781 ACTGTGACAGGTGACATGACATGACGCGTGTGTGTGTGATGATGATGATGATGATG 840
 Db 2334 ACTGTGACAGGTGACATGACATGACGCGTGTGTGTGTGATGATGATGATGATGATG 2393
 QY 841 ATCTTGATACCGGATGCTTCAAGCTGTGAGACCTTACAGGAGCATCTTCAATTCAG 900
 Db 2394 ATCTTGATACCGGATGCTTCAAGCTGTGAGACCTTACAGGAGCATCTTCAATTCAG 2453
 QY 901 CAAGCTATTTGAGCCACACAGAAACAGTACGCTGATGATGATGATGATGATGATGATG 960
 Db 2454 CAAGCTATTTGAGCCACACAGAAACAGTACGCTGATGATGATGATGATGATGATGATG 2513
 QY 961 AGCTATCATCTTACAGTGTCTTTGATGATCAACGCAAGTATGATGATGATGATGATGATG 1020
 Db 2514 AGCTATCATCTTACAGTGTCTTTGATGATCAACGCAAGTATGATGATGATGATGATGATG 2573
 QY 1021 GCTTATACAGCAGAGATCAAGGCTTCTGACAGTGTGATGATGATGATGATGATGATGATG 1080
 Db 2574 GCTTATACAGCAGAGATCAAGGCTTCTGACAGTGTGATGATGATGATGATGATGATGATG 2633

QY 1081 CAGAAATGATCTTGGGAGATGTGTCTTCTGATGATCAACAGGCTTCTTGAACAGGCC 1140
 Db 2634 CAGAAATGATCTTGGGAGATGTGTCTTCTGATGATCAACAGGCTTCTTGAACAGGCC 2693
 QY 1141 AACCACTCTGTTGGCTTATGCTTAAAGCATCTGCA 1173
 Db 2694 AACCACTCTGTTGGCTTATGCTTAAAGCATCTGCA 2726
 RESULT 3
 BOVCHYMOSIN 1305 bp mRNA linear MAM 26-APR-1993
 LOCUS bovine chymosin b (rennin) mRNA.
 DEFINITION J00003
 ACCESSION J00003.1 GI:162859
 VERSION chymosin; chymosin B; protease; rennin.
 KEYWORDS SOURCE
 ORGANISM Bos taurus (cow)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
 1 (bases 1 to 1305)
 HARRIS,T.J., LOWE,P.A., LYONS,A., THOMAS,P.G., EATON,M.A., MILLICAN,T.A., PATEL,T.P., BOSE,C.C., CAREY,N.H. and DOEL,M.T.
 Molecular cloning and nucleotide sequence of cDNA coding for calf preprochymosin
 Nucleic Acids Res. 10 (7), 2177-2187 (1982)
 JOURNAL MEDLINE PUBMED
 COMMENT Original source text: bovine (calf) cdna of fourth stomach mucosa mRNA.
 chymosin is the major proteolytic enzyme in the fourth stomach of the unweaned calf. two chromatographically different forms, a and b, of the enzyme and its precursor are known and a third form seems likely (see bovyhymoa, bovyhymoc). this sequence has tentatively been identified as coding for preprochymosin b. sequence comparison indicates that the precursors for chymosins a and b differ by only two amino acids, and for b and c by only four amino acids.
 FEATURES
 source location/Qualifiers
 1..1305
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 <1..1305
 /product="chymosin mRNA"
 26..1171
 /note="preprochymosin b"
 /codon_start=1
 /protein_id="AAA30448.1"
 /db_xref="GI:162860"
 /translation="MRCIVLLAVFALSGARIFRIPLYKSKSLRKALKERHLEDFLQKQDGISSKSYSGFEVAVSPVLTNYLDSQYFGKYLTPPEFTVLPFGSSDFWPSIYCKSNACKNORPDRKSTFONLGRPLSHYGSGNOGLIGDVTYVSNIVDIQOTVGLSTOBPDVFTVAEPDGLIGMAVPSLASVSLIPFNNMNRHLVADLFSVMDRNGOESMLTGLAINPSYTGSLHNPVPTVOOYMOFTSIVTISGVVAVAGCGCOALIDTGTSLVGPSSDIINTQOATGATONOGEDIDCDNLSNVPYFETENGMYPLTPSAVTSODGFCITSGFOSRNSQKILGDVFIIRYISVPRANNIVGLAKAT"
 sig_peptide 29..73
 /note="presquence"
 mat_peptide 200..1168
 /product="chymosin b"
 BASE COUNT 305 a 393 c 340 g 267 t
 ORIGIN psr-1 site.
 Query Match 76.8%; Score 900.4; DB 4; Length 1305;
 Best Local Similarity 87.8%; Pred. No. 1.3e-265;
 Matches 982; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
 QY 56 TCGTGTCTTACTACAGCTGTGATGATCAACCGATTCCTTCTCAACAAAGTATGCTC 115
 Db 54 TCTTGCTCTCTCCCAAGCGCTGATGATCACAGATCCCTCTGTACAAAGGCAAGTCTC 113
 QY 116 TCCGTAAAGCGCTGAAGAAATGAGACTTCTGAGAGACTTCTTGCAGAAACAAGATAG 175

```

Db      114 TGAAGAGGCGCTGAGAGACATGGCTTCTGAGGACCTTCGACAGAAACAGCAGTATG 173
QY      176 GCATCAGCAGCAAGTACTCCGGCTTCGGTGAAGTTGCTAGCGTCCACTTACCAATACC 235
Db      174 GCATCAGCAGCAAGTACTCCGGCTTCGGGAGGAGGCGCAGCGTGGCCCTGACCAATAC 233
QY      236 TTGATATGCAATCTTTGGGAGAGTCTACTCCGAAACCCCGCTCAAGAGTTGACCGTTTC 295
Db      234 TGAATAGTCAGTACTTTGGGAGAAATCACTCGGAGCCCGCCAGAGATTTACCGTGC 293
QY      296 TCTTTGATAGTCTGCTCTGACTTCTGAGGTTCCCTCTATCTACTGCAAGAGCAATGCT 355
Db      294 TGTTTGACATGGCTCTCTGACTTCTGGGTAACCTCTATCTACTGCAAGAGCAATGCT 353
QY      356 GCAAGAACCAACCAAGATTGATCCGAGAAAGTGTCCACCTTCCAGAACTTAGGCAAAC 415
Db      354 GCAAAAACCAACGCGCTCGACCCGAGAAAGTGTCCACCTTCCAGAACTTAGGCAAGC 413
QY      416 CCTTGTCTATACACTAGGTAAGGTAAGTACAGAGAACTTTAGGCTATGATACCGTCA 475
Db      414 CCTGTCTATACCACTAGGGAAGGAGCATGACGGGCATCTTAGGCTATGACACCGTCA 473
QY      476 CTGTCTCAACATTTGGAATTCACAGACAGTAGGACTTAGACCCCAAGAACAGGTG 535
Db      474 CTGTCTCAACATTTGGAATTCAGACAGAGCATGAGGCTGAGCAGCCAGAGAGCGGG 533
QY      536 ATGTCTTCACTTATGCAAAATTCGATGGCATCTTGGTATGGCAATCCATCGCTCGCT 595
Db      534 ACCTCTTCACTTATGCGAAATTCGACGGGATTCGAGGATGCGCTTACCCCTCGCTCGCT 593
QY      596 CAGAGTATGATACCTGTGTGTAACAATGATGAACCGACACTTATGATGCTCAAGACT 655
Db      594 CAGAGTATGATACCTGTGTGTAACAATGATGAACAGGCACTGTGTGCGCCCAAGACC 653
QY      656 TGTCTCGGTTTACATGAGCAGGAATGGCCAGAGAGACATGCTCAACCTTGGAGCTATTG 715
Db      654 TGTCTCGGTTTACATGAGCAGGAATGGCCAGAGAGACATGCTCAACCTTGGAGGCTATCA 713
QY      716 ATCCATCTTACCTACAGAGATCTCTTCACTGGGTTCCAGTCACTGTGACAGACATGAGC 775
Db      714 ACCGCTCTACTACAGAGGCTCCCTGCACTGGGTCGCGTGAAGTGAAGTGAAGTACGAGC 773
QY      776 AATTCACTGTGACAGGTGATCACTACGCGGTGTGTGTTGATGATGAAAGTGTGATGTC 835
Db      774 AGTTCACTGTGACAGGTGATCACTACGCGGTGTGTGTTGATGATGAAAGTGTGATGTC 833
QY      836 AAGCTATCTTGAATACCGGTACGTCAGAGCTGTGCGAACCTAGCAGGACATCTCAACA 895
Db      834 AAGCCATCTTGAACCGGCACTCTCAAGCTGTGTGGGCCCGAGCAGGACATCTCAACA 893
QY      896 TTCAGCAAGCTATTGAGCAACAGAACAGTACGCTGATGTTGACATGATTTGCGACA 955
Db      894 TTCAGCAGGCGCATTTGAGCCACAGAACAGTACGCTGATGTTGACATGATTTGCGACA 953
QY      956 ACCTTACGTAAGTCTGCTTCACTGTTGTCTTTGATGACAAACGAGATGATCCACTGACC 1015
Db      954 ACCTTACGTAAGTCTGCTTCACTGTTGTCTTTGATGACAAATGCGAAATGTACCCACTGACC 1013
QY      1016 CCTCCGCTTACAGCAGGAGATCAAGGTTCTGACACAGTGAATTCAGAGTGAAGAAC 1075
Db      1014 CCTCCGCTTACAGCAGGAGATCAAGGCTTCTGTACAGTGAATTCAGAGTGAAGAAC 1073
QY      1076 AATCCCAAGAAATGATCTTGGAGATGTGTATCTGTGATCTACAGGCTTCTTGAACA 1135
Db      1074 AATCCCAAGAAATGATCTTGGAGATGTGTATCTGTGATCTACAGGCTTCTTGAACA 1133
QY      1136 GGGCCAAACAACTGTGTGGGCTAGCTAAGCAATCTGA 1173
Db      1134 GGGCCAAACAACTGTGTGGGCTAGCTAAGCAATCTGA 1171

```

RESULT 4

```

AR002347
LOCUS      AR002347      1240 bp      DNA      linear      PAT 04-DEC-1998
DEFINITION Sequence 2 from patent US 5741665.
ACCESSION AR002347
VERSION    AR002347.1 GI:3963901
KEYWORDS   .
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   Unclassified.
AUTHORS     1 (bases 1 to 1240)
TITLE        Kato, E.K. and Stuart, M. Dorsey.
             Light-regulated promoters for production of heterologous proteins
             in filamentous fungi.
JOURNAL      Patent: US 5741665-A 2 21-APR-1998;
FEATURES     Location/Qualifiers
             source          1..1240
BASE COUNT   274 a      374 c      339 g      253 t
ORIGIN
Query Match 76.5%; Score 897.2; DB 6; Length 1240;
Best Local Similarity 87.7%; Pred. No. 1.3e-264;
Matches 980; Conservative 0; Mismatches 138; Indels 0; Gaps 0;
QY      56 TCGTGTCTTACTACACGCTGCTGAGATGACCCGCAATTCCTCTTCAAGGTAAGTCTC 115
Db      100 TCTTGTCTCTCTCCAGGGGCGTGAATGACAGAGATCCCTCTGTACAAAGGCAAGTCTC 159
QY      116 TCCGTAAGGCGCTGAAGAAACATGATGACTTGTAGAAACCTTTGCAAGAACACAGATAG 175
Db      160 TGAAGAGGCGCTGAAGAGAGCATGGCTCTGTGAGGACATTCCTCGAGAAACAGCAGTATG 219
QY      176 GCATCAGCAGCAAGTACTCCGGCTTCGGTGAAGTTGCTAGCGTCCACTTACCACTTACC 235
Db      220 GCATCAGCAGCAAGTACTCCGGCTTCGGGAGGAGTGCCAGCGTCCCTGACCAACTTACC 279
QY      236 TTGATATGCAATCTTTGGGAGAGATCTACTCGGAAACCCCGCTCAAGATTCACCGTTTC 295
Db      280 TGAATAGTCAGTACTTTGGGAGAGATCTACTCGGAAACCCCGCTCAAGATTCACCGTTTC 339
QY      296 TCTTTGATATCGGTTCTCTGACTTCTGTGGGTTCCCTCTATCTACTGCAAGAGCAATGCTC 355
Db      340 TGTTTGACATGGCTCTCTGACTTCTGTGGTACCTCTATCTACTGCAAGAGCAATGCTC 399
QY      356 GCAAGAACCAACCAAGATTGATTCGAGAAAGTGTGCACTTTCAGAACTTAGGCAAAC 415
Db      400 GCAAAAACCAACGAGCTTGCAGCCGAGAAAGTGTGCACTTTCAGAACTTAGGCAAGC 459
QY      416 CCTTGTCTATACCTACGTAAGGTAAGTGAATGTAAGAAATCTTAGGCTATGATACCGTCA 475
Db      460 CCTGTCTATACCTACGTAAGGTAAGTGAATGTAAGAAATCTTAGGCTATGATACCGTCA 519
QY      476 CTGTCTCAACATTTGAGCAATTCACAGACAGTGAAGATTCAGCAACCCCAAGAACAGGTG 535
Db      520 CTGTCTCAACATTTGAGCAATTCAGAGACATGAGAGCAATGAGCTTGAACCCAGAGACC 579
QY      536 ATGTCTTCACTTATGCAAAATTCGATGGCATCTTGGTATGAGTAATCCATGCTGTGCGT 595
Db      580 ACGTCTTCACTTATGCAAAATTCGAGGAGTCTTGGGAGTGGCTTACCTCGCTGCGCT 639
QY      596 CAGAGTATCTGATACCTGTGTGTAACAATGATGAACCGACACTTATGATGCTCAAGACT 655
Db      640 CAGAGTATCTGATACCTGTGTGTAACAATGATGAACCGACACTTATGATGCTCAAGACT 699
QY      656 TGTCTCGGTTTACATGAGCAGGAATGGCCAGAGAGACATGCTCAGGCTTGAAGTATTG 715
Db      700 TGTCTCGGTTTACATGAGCAGGAATGGCCAGAGAGACATGCTCAGGCTTGAAGTATTG 759
QY      716 ATCCATCTTACATACAGAGATCTTCACTGGGTTCAAGTCACTGTGACAGCACTAGTGC 775
Db      760 ACCGCTCTACTACACAGAGGCTCCCTGCACTGGGTGCGCGTGAAGTGAAGTGAATGTC 819
QY      776 AATTCACTGTGACAGGTGATCACTCAGCGGCTGTGTTGTGATGATGAAGTGAATGTC 835

```

	REFERENCE	TITLE	JOURNAL	COMMENT
Db	820	AGTTCACTGTGGAAGAATGTACACTACAGCGGTGTGGTTCCTGTGAGSGTGCTGTC	879	
OY	836	AAGCTATTCTTGGATACC GGTAACGTCCAGCTGTGTGCACCTTAGCAGCCGACATTTCTACA	895	
Db	880	AGGCCATCTCTGGACACGGGACACCTCCAAGCTGTGTGGGCCACAGCAGGACCATCTCAACA	939	
OY	896	TTCCAGCAAGCTATTGGAGCCACACAGAACCCAGTACCGGTGAATTGGACATAAGTTGGCACA	955	
Db	940	TCGACAGAGCCATTGGAGCCACACAGAACCCAGTACCGGTGAATTGGACATAAGTTGGCACA	999	
OY	956	ACCTTAGCTACATGCGCTACAGTTGCTTTGGATCAACGGCAGAGTATACCACCTGACCC	1015	
Db	1000	ACCTGAGCTACATGCCACTGTGCTTTGGATCAATGGCAAATGTATCCCACTGACCC	1059	
OY	1016	CCTCCGCTTATACAGACCCAGGATCAAGGGTTCTGCACCACTGATATTCAGAGTGAAC	1075	
Db	1060	CCTCCGCTTATACAGACCCAGGATCAAGGGTTCTGCACCACTGATATTCAGAGTGAAC	1119	
OY	1076	ATTCCCAGAAATGGATCTTTGGAGATGTGTTTCATTCGTGAGTACTACAGGCTCTTTGACA	1135	
Db	1120	ATTCCCAGAAATGGATCTTTGGAGATGTGTTTCATTCGTGAGTACTACAGGCTCTTTGACA	1179	
OY	1136	GGGCCACACACCTGCTGGTGGGCTAGCTAAGCAATCTGA	1173	
Db	1180	GGGCCACACACCTGCTGGGCTAGCTAAGCAATCTGA	1217	
RESULT 5	E00042	E00042	1460 bp	RNA linear PAT 29-SEP-1997
LOCUS	DEFINITION	DNA coding of pre-pro rennin.		
VERSION	ACCESSION	E00042.1	GI:2168348	
KEYWORDS	KEYWORDS	JP 1982141287-A/1.		
SOURCE	SOURCE	unidentified		
ORGANISM	ORGANISM	unidentified		
REFERENCE	REFERENCE	1 (bases 1 to 1460)		
AUTHORS	AUTHORS	Baanadeste,R.A., Jieni,M., Donarudo,T.M., Arison,T.R. and Jierardo,F.B.		
TITLE	TITLE	RENNIN, PREPRORENNIN OR PRORENNIN GENE OBTAINED FROM RECOMBINED DNA MATERIAL AND LIVE CELL CONTAINING GENE Patent: JP 1982141287-A 1 01-SEP-1982;		
JOURNAL	JOURNAL	KORABORATEIBU RES INC		
COMMENT	COMMENT	OS calf PN JP 1982141287-A/1 PD 01-SEP-1982 JP 1982003556 PR 16-JAN-1982 US 81 225717, 01-DEC-1981 US 81 325481 PI BANAADESTE RABUTSUTUKI ARUFUODOO, JIENI MAO, PI DONARUDO TEIRAA MOIAA. PI ARISON TANDUNTON RIGUBII, JIERARDO FURANSISHU BUOBUSU PC C1ZNI/00,C0767/00,C07H21/04,C1ZNI5/00,C12P21/00//C1ZNI/18, PC C1ZNI/20, PC C12P19/34,C12R1/19,C12R1/865; CC strandedness: Double; CC topology: Linear; CC hypothetical: No; CC anti-sense: No; CC *source: tissue Location/Qualifiers FH Key FT CDS 205..1350 /product='pre-pro rennin' FT mat_peptide 205..1347 /product='pre-pro rennin' FT CDS 253..1350 /product='pro rennin' FT mat_peptide 253..1347 /product='pro rennin' FT CDS 278..1350 /product='rennin'		

FEATURES	FT	mat_peptide	278..1347
SOURCE	FT	Location/Qualifiers	/product="remind"
		1..1460	/organism="unidentified"
		/mol_type="genomic RNA"	/db_xref="taxon:32644"
BASE COUNT	328 a	440 c	398 g
ORIGIN			294 t
Query Match	76.5%;	Score 897.2;	DB 6; Length 1460;
Best Local Similarity	87.7%;	Pred. No. 1.3e-264;	
Matches	980; Conservative	0; Mismatches 138; Indels	0; Gaps 0;
QY	56	TCGTGCTGTTACTACGCGCTGTGAGATCAACCGCATTCCTCTCTACAAAGTAAGTCTC	115
Db	233	TCCTGCTCTCTCCAGAGGCGGTGATGATCAACAGAGATCCCTCTGTACAAAGGCAAGTCTC	292
QY	116	TCGCTAAGGGGCTGAAAGAACATGAGATCTTCTAGAAAGCTTCTGCAAAACACAGATATG	175
Db	293	TGAGGAAGGGCGCTGAAAGAGCATGGGCTTCTGAGAGACTTCTGCAAGAAACAGAGATATG	352
QY	176	GCATCAGCAGCAAGTATCTCCGGCTTCGGTGAAGTGTCTAGCGTCCACTTACCAACTATCC	235
Db	353	GCATCAGCAGCAAGTATCTCCGGCTTCGGGAGGTGTGCGACAGGTGCCCCCTGACCAACTATCC	412
QY	236	TTGATAGTCAATATCTTTGGAGAGATCTACTCTCGAACCCCGCTCAAGAGTTCAACGGTTC	295
Db	413	TGATAGTCAAGTATCTTTGGAGAGATCTACTCTCGGAGCCCCGCCAGAGAGTTCAACCGTTC	472
QY	296	TCCTTGAATCTGGTTCCTCTGACTTCTGCGGTTCCCTCTATCTACTGCAAGAGCAATGCTT	355
Db	473	TGTTTGAATCTGGTTCCTCTGACTTCTGCGGATCCCTCTATCTACTGCAAGAGCAATGCTT	532
QY	356	GCAAGAACCAACCAAGATTCGATCCGAGAAAGTGTGTCACCTTCCAAAGCTTAGGCAAC	415
Db	533	GCAAGAACCAACCAAGCTTCGATCCGAGAAAGTGTGTCACCTTCCAAAGCTTAGGCAAC	592
QY	416	CCCTGCTATATCACTACGCTGATAGGTGATGCAAGAGATCTTTAGGCTATGATACCGTCA	475
Db	593	CCCTGCTATATCACTACGCTGATAGGTGATGCAAGAGATCTTTAGGCTATGATACCGTCA	652
QY	476	CTGTCTCAACATTTGGAATTTCAACAGACAGTATGAGCTTAGAGCAACCAAGAACAGAGTG	535
Db	653	CTGTCTCAACATTTGGAATTTCAACAGACAGTATGAGCTTAGAGCAACCAAGAACAGAGTG	712
QY	536	ATGCTCTCAACATTTGGAATTTCAACAGACAGTATGAGCTTTGATATGAGCTATCCATCCGCTCGCT	595
Db	713	ACGCTCTCAACATTTGGAATTTCAACAGACAGTATGAGCTTTGATATGAGCTATCCATCCGCTCGCT	772
QY	596	CAGAGTACTGATATCTCTGTGTTTGAACAACATGATGAAACCGACACTTATGATCTCAGACT	655
Db	773	CAGAGTACTGATATCTCTGTGTTTGAACAACATGATGAAACCGACACTTATGATCTCAGACT	832
QY	656	TGTTCTCGGTTTACATGAGACAGAAATGGCCAGAGAGACATGCTCACGCTTGAAGCTATG	715
Db	833	TGTTCTCGGTTTACATGAGACAGAAATGGCCAGAGAGACATGCTCACGCTTGAAGCTATG	892
QY	716	ATCATCTCTACTACACAGATCTCTTCACTGCGGTTCCAGTACCTGTGACAGATGACTGGC	775
Db	893	ACCGTCTCTACTACACAGAGGCTCTGTGACCTGGAGCCCTGTGACAGTCAAGAGTATCGGC	952
QY	776	AATTCACCTGAGGACAGGTGTCACCATCAAGCGGTTGTTGTCATGTGAAGGTGATGTC	835
Db	953	AGTTTCACCTGAGGACAGGTGTCACCATCAAGCGGTTGTTGTCATGTGAAGGTGATGTC	1012
QY	836	AAGCTATCTTGATATCAAGGTGTCACCATCAAGCTGTGAGCCTGACAGGACATTTCTCAAC	895
Db	1013	AGGCTATCTTGATATCAAGGTGTCACCATCAAGCTGTGAGCCTGACAGGACATTTCTCAAC	1072
QY	896	TTTCAAGCAAGCTATTTGAGGACACACAGAACCAAGTACGCTGATTTGAACATGATTTGGACAC	955
Db	1073	TTTCAAGCAAGCTATTTGAGGACACACAGAACCAAGTACGATGATGATTTGAACATGATTTGGACAC	1132

```
QY 956 ACCTTAGCTACATGCTTACAGTTGTCTTTGAGATCAACGGCAAGATGACCTGACCC 1015
DB 1133 ACCTGAGCTACAGATGCCACGTGTCTTTGAGATCAATGCGAAATGATCCCACTGACCC 1192
QY 1016 CCTCCGCTATACACAGCAGAGATCAAGGGTTCTGCACCACTGGATTCAGAGTGAGAAC 1075
DB 1193 CCTCCGCTATACAGCAGCAGCAGGCTTCTGTACCACTGGCTTCAAGATGAAATC 1252
QY 1076 ATTCGCAAGAAATGATCTTGGAGATGTGTTCATTCTGTAGTACTACAGCGTCTTGACA 1135
DB 1253 ATTCGCAAGAAATGATCTTGGAGATGTGTTCATTCTGTAGTACTACAGCGTCTTGACA 1112
QY 1136 GGGCCAAACAACCTTCGTGGGCTGAGTAAAGCAATCTGA 1173
DB 1313 GGGCCAAACAACCTTCGTGGGCTGAGTAAAGCAATCTGA 1350

RESULT 6
A15836
LOCUS A15836 1291 bp DNA linear PAT 16-FEB-1994
DEFINITION chymosin gene.
ACCESSION A15836
VERSION A15836.1 GI:488959
KEYWORDS
SOURCE
ORGANISM Bos taurus (cow)
            Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovine; Bos.
            1 (bases 1 to 1291)
REFERENCE
AUTHORS Simons,A.F.M. and De Vos,W.M.
TITLE DNA fragments, containing a lactic acid bacterium-specific
        regulator region for the expression of genes coding for normally
        heterologous proteins
        Patent: EP 0307011-A 5 15-MAR-1989;
        NEDERLANDS INSTITUUT VOOR ZUIVELONDERZOEK
        Location/Qualifiers
            1..1291
            /organism="Bos taurus"
            /mol_type="genomic DNA"
            /db_xref="taxon:9913"
BASE COUNT 323 a 382 c 328 g 257 t 1 others
ORIGIN
Query Match 76.4%; Score 896.2; DB 6; Length 1291;
Best Local Similarity 87.6%; Pred. No. 2,7e-264;
Matches 979; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 56 TCGTGTCTTACTACGCTGCTGAGATCACCGCATTCCTCTTACAAAGTAACTTC 115
DB 9 TCTTCGCTCTCTCCACAGGCGCTGAGATCACAGATCCCTCTGTACAAAGGCAAGTCTC 68

QY 116 TCGTAGAGGCGCTGAAGAAATGAGACTTCTAGAAAGCTTGGCAAAACAAGATAG 175
DB 69 TGAAGAAAGCGCTGAAGAGCATGAGGCTTCTGAGAGACTTCTGCAAAACAAGAGATAG 128
QY 176 GCATCAGCAAGATCTCCGCGCTTCGATGAAGTTGCTAGCGTCCACTTACCAACTACC 235
DB 129 GCATCAGCAAGATCTCCGCGCTTCGAGAGAGTGGCAGAGTGGCCCTTACCAACTACC 188
QY 236 TTGATAGTCAATATCTTTGGAGAGTCTACCTCGAACCCTCGCTCAAGAGTTCAACGTTT 295
DB 189 TGAATAGTCAATATCTTTGGAGAGTCTACCTCGAACCCTCGCTCAAGAGTTCAACGTTT 248
QY 296 TCTTTGATCTGTTCTCTGATCTTCTGGGTTCCCTCTATCTATCTAGCAAGCAATGCTT 355
DB 249 TGTTTGACCTGCTCTCTGATCTTCTGGGTTCCCTCTATCTATCTAGCAAGCAATGCTT 308
QY 356 GCAAGAACCAACAAAGATTCGATCCGAGAAAGTGTCCACTTCCAGAACTTAGCAAAAC 415
DB 309 GCAAGAACCAACAGCGCTTGAACCCGAGAAAGTGTCTCACTTCCAGAACTTAGCAAAAC 368
```

```
QY 416 CTTTGTCTATACACTACCGGTACAGAGTAGATGCAAGGAATCTTAGGCTATGATACCGTCA 475
DB 369 CCTGTCTATCCACTACGGAACAGGAGCATGAGGCAATCCTGGCTATGACCGTCA 428
QY 476 CTGTCTTCAACATTTGAGCATTTCAACAGACAGTAGAGACTTAGACCCCAAGAACAGGTG 535
DB 429 CTGTCTTCAACATTTGAGCATTTCAACAGACAGTAGAGACTTAGACCCCAAGAACAGGTG 488
QY 536 ATGTCTTCACTATGCAAAATTCGATGCGATCTTGTGTATGCGATTCCTTGGCTTGGCT 595
DB 489 ACGTCTTCACTATGCAAAATTCGATGCGATCTTGTGTATGCGATTCCTTGGCTTGGCT 548
QY 596 CAGAGTACTGATACCTGTTGTTGACATATGAAACCGAACCACTAGTACCTCAAGACT 655
DB 549 CAGAGTACTGATACCTGTTGTTGACATATGAAACCGAACCACTAGTACCTCAAGACT 608
QY 656 TGTTCCTCGTTTACATGACAGGAATGGCCAGAGAGCATGCTACGCTTGGAGTATTG 715
DB 609 TGTTCCTCGTTTACATGACAGGAATGGCCAGAGAGCATGCTACGCTTGGAGTATTG 668
QY 716 ATTCATCTTACTACACAGATCTTCTCACTGGGTTCCAGTCACTGTGCAGCACTATGCG 775
DB 669 ACCCTTCTTACTACACAGATCTTCTCACTGGGTTCCAGTCACTGTGTGAGCACTATGCG 728
QY 776 AATTCACTGTGACAGATGACCATATGAGGAGTGTGTTGTCATGTAAGGTGATGTC 835
DB 729 AGTTCACTGTGACAGATGACCATATGAGGAGTGTGTTGTCATGTAAGGTGATGTC 788
QY 836 AAGCTATCTTGTGATCCGATACGTCCAAAGCTGTGCGAATCTAGCAGCAGCATTTCAACA 895
DB 789 AAGCCATCTGTGACAGGAGCACTTCAAGCTGTGCGGCGCCAGCAGCATCTTCAACA 848
QY 896 TTGAGAAAGTATTTGAGCCACACAGAACCATGACGTGATTTGACATGATTTGCGACA 955
DB 849 TTCAGCAGGCGATTTGAGCCACACAGAACCATGACGTGATTTGACATGATTTGCGACA 908
QY 956 ACCTTAGCTACATGCTTACATGTTCTTTGAGATCAACGGCAAGATGACCTGACCC 1015
DB 909 ACCTTAGCTACATGCTTACATGTTCTTTGAGATCAATGCAAAATGATCCCACTGACCC 968
QY 1016 CCTCCGCTATACACAGCAGATCAAGGGTTCTGCACCACTGGATTCAGAGTGAGAAC 1075
DB 969 CCTCCGCTATACACAGCAGATCAAGGGTTCTGCACCACTGGATTCAGAGTGAGAAC 1028
QY 1076 ATTCGCAAGAAATGATCTTGGAGATGTGTTCATTCTGTAGTACTACAGCGTCTTGACA 1135
DB 1029 ATTCGCAAGAAATGATCTTGGAGATGTGTTCATTCTGTAGTACTACAGCGTCTTGACA 1088
QY 1136 GGGCCAAACAACCTTCGTGGGCTGAGTAAAGCAATCTGA 1173
DB 1089 GGGCCAAACAACCTTCGTGGGCTGAGTAAAGCAATCTGA 1126

RESULT 7
104058
LOCUS 104058 1175 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 5 from Patent EP 0123928.
ACCESSION 104058
VERSION 104058.1 GI:591912
KEYWORDS
SOURCE
ORGANISM Unknown.
            Unclabified.
            1 (bases 1 to 1175)
REFERENCE
AUTHORS Cashion,L.M., McManan,M.T., Rice,C.W. and Sias,S.R.
TITLE Recombinant DNA coding for a polypeptide displaying milk clotting
        activity
        Patent: EP 0123928-A2 5 07-NOV-1984;
        Location/Qualifiers
            1..1175
            /organism="unknown"
BASE COUNT 263 a 356 c 318 g 238 t
ORIGIN
```

Query Match 76.4%; Score 895.6; DB 6; Length 1175;
Best Local Similarity 87.6%; Pred. No. 4e-264;
Matches 979; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

```
QY 56 TCGTGTCTGTTACTACACGCTGTGATGATCAACCGCATTCCTCTTACAAAGGTAAGTCTC 115
DB 49 TCTTGTCTCTCTCCCGAGGCGCTGAGATCACAGGATCCCTCTGTACAAAGCAAGTCTC 108
QY 116 TCCGTAAGCGCTGAAGAACTGACCTTCTAGAGACTCTCTGAGAAACAAAGCATATG 175
DB 109 TGAGGAAGCGCTGAGAGAGCATGAGGCTCTGAGAGACTCTCTGAGAAACAGCATATG 168
QY 176 GCATCAGCAGCAAGTACTCCGCTTCGAGTGAAGTGTCTAGCTGACCTTACCACTAC 235
DB 169 GCATCAGCAGCAAGTACTCCGCTTCGAGAGAGTGGCCAGCTGCCCCCTGACCAACTAC 228
QY 236 TTGATGATCAATCTTTGAGAAAGATCTACCTCGAAACCCCGCTCAAGATTCACCGTTC 295
DB 229 TGGATAGTCAAGTACTTTGAGAAAGATCTACCTCGGACCCCGCCAGAGTTGACCGTTC 288
QY 296 TCTTGAATAGTGTCTCTCTGATCTTCTGAGTCTCTCTATCTACTGCAAGAGCATGCT 355
DB 289 TGTGTGACACTGCTCTCTGATCTTCTGAGTACCTCTATCTACTGCAAGAGCATGCTCT 348
QY 356 GCAAGAAACCAAGAAAGTTCGATCCGAAAGTCTGCACTTCAGAACTTGAAGCAAC 415
DB 349 GCAAAACCAAGAGGCTTTCAGACCCGAAAGTCTCTCACTTCAGAACTTGAAGCAAC 408
QY 416 CTTGTCTATACACTACGATACAGGTACATGCAAGAACTTGAAGTATGATACCTGCA 475
DB 409 CCCTGTCTATCCACTACAGGAGCAGCAGCATGCAAGGAGCTCTAGAGTATGACAGCTCA 468
QY 476 CTGTCTCAAACTTGTGACATCCAGAGACAGTGGCTGAGCAACCAAGCAAGTGTG 535
DB 469 CTGTCTCAAACTTGTGACATCCAGAGACAGTGGCTGAGCAACCAAGCAAGTGTG 528
QY 536 ATGTCTTACCTATGACAGAAATTCAGTGGATCTCTTGTATGATACCATGCTCGAGT 595
DB 529 ACGTCTTACCTATGACAGAAATTCAGTGGATCTCTTGTATGATACCATGCTCGAGT 588
QY 596 CAGAGTACTGATACCTGTGTGTTGACAACTATGATGAACCGACCTAGTACTGCAAGT 655
DB 589 CAGAGTACTGATACCTGTGTGTTGACAACTATGATGAACCGACCTAGTACTGCAAGT 648
QY 656 TGTCTCGGTTTACATGACAGAGATGGCCAGAGAGCATGCTACGCTTGAAGCTATG 715
DB 649 TGTCTCGGTTTACATGACAGAGATGGCCAGAGAGCATGCTACGCTTGAAGCTATG 708
QY 716 ATTCATCTACTACACAGAGATCTCTTCACTGGGTTCACTGTCAGAGCATGCTGGC 775
DB 709 ACCGCTCTACTACACAGAGATCTCTTCACTGGGTTCCGTCAGATGTCAGAGTCTGGC 768
QY 776 AATTCATCTGTCAGATGTCACCATCAGCGGTGTGTGTTGTCATGTGAGAGTGTGATGTC 835
DB 769 AGTTCATCTGTCAGATGTCACCATCAGCGGTGTGTGTTGTCATGTGAGAGTGTGATGTC 828
QY 836 AAGCTATCTTGGATACCGGATGCTGCAAGCTGTGCACTGACAGCACTTCTCAACA 895
DB 829 AGGCTATCTTGGATACCGGATGCTGCAAGCTGTGCAAGCTGTGCAAGCTTCTCAACA 888
QY 896 TTTCAGCAGCTATTTGAGAGCACAAGAACCAAGTACAGGATGTTACATATGTCGACA 955
DB 889 TCCAGCAGCTATTTGAGAGCACAAGAACCAAGTACAGGATGTTACATATGTCGACA 948
QY 956 ACCTTAGTACATGCTCAAGTGTCTTTGAGATCAACGGCAAGATGTAACCACTGACCC 1015
DB 949 ACCTTAGTACATGCTCAAGTGTCTTTGAGATCAACGGCAAGATGTAACCACTGACCC 1008
QY 1016 CCTCGGCTATACAGGAGATCAAGGTTTGTGACCAAGTGTGATTTCCAGAGTGAAGC 1075
DB 1009 CCTCGGCTATACAGGAGATCAAGGTTTGTGACCAAGTGTGATTTCCAGAGTGAAGTTC 1068
```

```
QY 1076 ATTCGCAAAATGATCTTGGAGATGTGTTCAATTCGTGAGTACTACAGCGTCTTTGACA 1135
DB 1069 ATTCGCAAAATGATCTTGGAGATGTGTTCAATTCGTGAGTACTACAGCGTCTTTGACA 1128
QY 1136 GGGCCAAACCTGTTGGGCTAGTAAAGCAATGTA 1173
DB 1129 GGGCCAAACCTGTTGGGCTAGTAAAGCAATGTA 1166
```

RESULT 8

AR073077

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

BASE COUNT

ORIGIN

Query Match 76.3%; Score 894.8; DB 6; Length 2733;
Best Local Similarity 88.4%; Pred. No. 7.6e-264;
Matches 971; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

```
QY 76 GCTGATGATACCCCGATTCCTCTTACAAAGTATGCTCTCCGTAAGGCGCTGAAGAA 135
DB 1630 GCTGATGATACCCCGATTCCTCTTACAAAGTATGCTCTCCGTAAGGCGCTGAAGAA 1689
QY 136 CATGACCTTCTAGAAAGATCTTCTGAGAAACAAAGATGAGATGAGCAGCAAGTATCC 195
DB 1690 CATGACCTTCTAGAAAGATCTTCTGAGAAACAAAGATGAGATGAGCAGCAAGTATCC 1749
QY 196 GCGTTCGAGTGAAGTGTCTAGGCGTGCACCTTACCACTACCTTGTATGATCAATATCTTGGG 255
DB 1750 GCGTTCGAGTGAAGTGTCTAGGCGTGCACCTTACCACTACCTTGTATGATCAATATCTTGGG 1809
QY 256 AAGATCTACCTGAGAAACCCCGCTTAAAGATTCACCGTCTCTTGTATGATGATGCTCT 315
DB 1810 AAGATCTACCTGAGAAACCCCGCTTAAAGATTCACCGTCTCTTGTATGATGATGCTCT 1869
QY 316 GACTCTGAGGTTCCCTCTATGATGATGAGCAAGTATGCTGCAAGAACCAAGAAATTC 375
DB 1870 GACTCTGAGGTTCCCTCTATGATGATGAGCAAGTATGCTGCAAGAACCAAGAAATTC 1929
QY 376 GATCCGAAAGTGTCTGACCTTCCAGAACTTGAAGCAACCTTGTATGATGATGATGATG 435
DB 1930 GATCCGAAAGTGTCTGACCTTCCAGAACTTGAAGCAACCTTGTATGATGATGATGATG 1989
QY 436 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 495
DB 1990 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2049
QY 496 ATTCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 555
DB 2050 ATTCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2109
QY 556 TTTCAGTGCATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 615
DB 2110 TTTCAGTGCATCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2169
QY 616 TTTCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 675
DB 2170 TTTCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2229
```


QY 836 AAGTATCTTGATACCGGATAGTCCAAAGCTGGTGGACCTAGCAGCCGACATTTCAACA 895
DB 829 AGGCATCTGGACACGGGACACCTCCAGCTGTGGGGCCGACGACCATCTTCAACA 888
QY 886 TTCAGCAAGCTATTGGAGCCACACAGAACCAAGTACGGTAGTTGATCATAGATTGCGACA 955
DB 889 TCCAGCAGGCGCATTTGGAGCCACACAGAACCAAGTACGATGATTTGATCATGCTCGACA 948
QY 956 ACCTAGTACATGCTACAGTGTCTTGGATGATCAAGCGCAAGATGATCCACTGACCC 1015
DB 949 ACCTAGTACATGCTACAGTGTCTTGGATGATCAAGCGCAAGATGATCCACTGACCC 1008
QY 1016 CCTCGCCTTATACAGCAGGATCAAGGATCTGACAGCAGTGGATTCCAGAGTGAACC 1075
DB 1009 CCTCGCCTTATACAGCAGGATCAAGGATCTGACAGCAGTGGATTCCAGAGTGAACC 1068
QY 1076 ATTCGCAAGATGATCTTGGAGATGTTGATTCATTCGTAGTACTACAGCGTCTTTGACA 1135
DB 1069 ATTCGCAAGATGATCTTGGAGATGTTGATTCATTCGTAGTACTACAGCGTCTTTGACA 1128
QY 1136 GGGCCAAACACTCTTGGGCTAGCTAAAGCAATCTGA 1173
DB 1129 GGGCCAAACACTCTTGGGCTAGCTAAAGCAATCTGA 1166

RESULT 10
LOCUS E00075 1290 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding chymosin (=rennin).
ACCESSION E00075
VERSION E00075.1 GI:2168379
KEYWORDS JP 1983009687-A/1.
SOURCE Bos taurus (cow).
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 1290)
Nomenclature: H.K., Maikere, T.D., Chimoshi, J.R.H.H., Pilita, A.R. and
Jiyon, S.E.
PRODUCTION OF POLYPEPTIDE
Patent: JP 1983009687-A 1 20-JAN-1983;
CELLTECH LTD
OS bovine
PN JP 1983009687-A/1
PD 20-JAN-1983
PR 17-JUN-1982 JP 1982104672
PR 17-JUN-1981 GB 81 8118688, 11-NOV-1981 GB 81 8133998, PR
01-DEC-1981 GB 81 8136185, 10-FEB-1982 GB 82 8203907 PI
NOOMAN HERRI KEARI, MAIKERU TERENSU DOORU,
PI CHIMOSHI JIYON KOI HARISU,
PI PILITA ANSONIT ROU, JIYON SUPENSAA EMUTBEJI
PC C12N9/52,C07H21/04,C12N1/00,C12N9/60,C12N15/00, PC
C12Q1/02//C12R1/19,
PC C12R1/865;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC *source: clone=pCT 70;
CC *source: clone=pCT 70;
FH key Location/Qualifiers
FH 5'UTR 1..25
FT sig_peptide 26..73
FT mat_peptide 74..1078
FT /product= 'pro-chymosin'
FT 3'UTR 1082..1200
FT CDS 26..1081
FT /product= 'pre-pro-chymosin'.
FEATURES
source Location/Qualifiers
1..1290
/organism= "Bos taurus"

BASE COUNT 309 a 384 c 334 g 263 t
ORIGIN
Query Match 76.2%; Score 894; DB 6; Length 1290;
Best Local Similarity 87.5%; Pred. No. 1.3e-263;
Matches 978; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
/mol_type="genomic RNA"
/db_xref="taxon:9913"
QY 56 TCGTGTCTGTTACTACGCTGCTGAGATCACCCGATTCCTCTTACAAAGTAACTGTC 115
DB 54 TCTTGTCTCTCTCCCAAGGCGCTGATCACAGATCCCTCTTACAAAGGCAAGTCTC 113
QY 116 TCCGTAAGCGCGTGAAGAACATGAGACTCTTGAAGACTCTTGGCAGAAACAAGTATG 175
DB 114 TGAGGAAGGCGGTGAGAGAGATGGGCTTCTGAGAGACTTCTGAGAAACAGCATATG 173
QY 176 GCATCAGCAGCAAGTACTCCGGCTTGGTGAAGTTGCTAGCGTGCACCTTACCAACTACC 235
DB 174 GCATCAGCAGCAAGTACTCCGGCTTGGTGAAGTTGCTAGCGTGCACCTTACCAACTACC 233
QY 236 TTGATAGTCAATCTTTGGAAAGATCTACCTCGAACCCTCGTCAAGATTCAACGTTTC 295
DB 234 TAGATAGTCAATCTTTGGAAAGATCTACCTCGAACCCTCGTCAAGATTCAACGTTTC 293
QY 296 TCTTTGATACGTTTCTCTGACTCTTGGGTTCCCTTATGACTGCAAGAGCAATGCT 355
DB 294 TGTTTGACACTGGCTCTCTGACTCTTGGGTTCCCTTATGACTGCAAGAGCAATGCT 353
QY 356 GCAAGAACCAACCAAGATTCGATCCGAAAGTCTGCCACCTTCCAGAACTTAGGCGCAAC 415
DB 354 GCAAGAACCAACCAAGATTCGATCCGAAAGTCTGCCACCTTCCAGAACTTAGGCGCAAC 413
QY 416 CTTGTCTTATACACTGCTGACAGGTACGATGACAGAACTTTAGGCTATGATACGTTCA 475
DB 414 CCTGTCTTATACACTGCTGACAGGTACGATGACAGAACTTTAGGCTATGATACGTTCA 473
QY 476 CTGTCTTCAAGATTTGTGACACCTTCAAGCAGATGACCTTACGCCAAGAACAGAGTG 535
DB 474 CTGTCTTCAAGATTTGTGACACCTTCAAGCAGATGACCTTACGCCAAGAACAGAGTG 533
QY 536 ATGTCTTCACTATGACAAATTCGATGAGCATCTTGTGATGACATACCATGCTGCGCT 595
DB 534 ATGTCTTCACTATGACAAATTCGATGAGCATCTTGTGATGACATACCATGCTGCGCT 593
QY 596 CAGAGTACTGATACCTGTGTTGACAAATGATGACCACTAGTACTGCAAGACT 655
DB 594 CAGAGTACTGATACCTGTGTTGACAAATGATGACCACTAGTACTGCAAGACT 653
QY 656 TGTTCCTCGTTTACATGACAGGAATGGCCAGAGAGATGTCACGCTTGAAGCTATTTG 715
DB 654 TGTTCCTCGTTTACATGACAGGAATGGCCAGAGAGATGTCACGCTTGAAGCTATTTG 713
QY 716 ATCACTCTTACTACACAGAGATCTCTTCACTGGGTTCCAGTCACTGTGACAGTACTGGC 775
DB 714 ACCGCTCTTACTACACAGAGATCTCTTCACTGGGTTCCAGTCACTGTGACAGTACTGGC 773
QY 776 AATTCATGTGACAGTGTCAACATCAGCGGTGTGTTGTCATGTGAAGGTGATGTC 835
DB 774 AATTCATGTGACAGTGTCAACATCAGCGGTGTGTTGTCATGTGAAGGTGATGTC 833
QY 836 AAGCTATCTTGGATACCGGTACAGTGTGTCAGAGCTTGAAGCGCACTTCTCAACA 895
DB 834 AAGCTATCTTGGATACCGGTACAGTGTGTCAGAGCTTGAAGCGCACTTCTCAACA 893
QY 896 TTCAGCAAGCTATTGGAGCCACACAGAACCAAGTACGGTAGTTGATCATAGATTGCGACA 955
DB 894 TCCAGCAGGCGCATTTGGAGCCACACAGAACCAAGTACGGTAGTTGATCATAGATTGCGACA 953
QY 956 ACCTAGTACATGCTACAGTGTCTTGGATGATCAAGCGCAAGATGATCCACTGACCC 1015
DB 954 ACCTAGTACATGCTACAGTGTCTTGGATGATCAAGCGCAAGATGATCCACTGACCC 1013


```

QY 1016 CCTCCGCTATACGACGATCAAGGCTTCTGACCAAGTGAATTCAGAGTGAAC 1075
DB 1014 CCTCCGCTATACGACGATCAAGGCTTCTGACCAAGTGAATTCAGAGTGAAC 1073
QY 1076 ATTCGCAAGAAATGATCTGGAGATGTGTTCAATGCTGATCTACAGGCTTTGACA 1135
DB 1074 ATTCGCAAGAAATGATCTGGAGATGTGTTCAATGCTGATCTACAGGCTTTGACA 1133
QY 1136 GGGCCACAACCTCTGTTGGCTGCTAAGCAATCTGA 1173
DB 1134 GGGCCACAACCTCTGTTGGCTGCTAAGCAATCTGA 1171

RESULT 11
A15633 1291 bp mRNA linear PAT 18-FEB-1994
LOCUS preprochymosin.
DEFINITION A15633
ACCESSION A15633
VERSION A15633.1 GI:491951
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1291)
AUTHORS Carey,N.H., Doel,M.T., Harris,T.J.R., Lowe,P.A. and Emtage,J.S.
TITLE A process for the production of a polypeptide
JOURNAL Patent: EP 0068691-A 29 05-JAN-1983;
CELLTECH LIMITED
FEATURES
source 1. 1291
location/Qualifiers
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
1. 1291
/gene="preprochymosin"
26. 1171
/gene="preprochymosin"
/codon_start=1
/transl_table=1
/protein_id="CAA01244.1"
/db_xref="GI:491952"
/translation="MRCLIVLTALVFPALSGAEITRIPLYKSLRLAKHGLLEDEL
OKOQVIGISKYSRGFGEVAVPLNVLDSQFQKITYLCTPROEPTVLPDSSSPFWPS
IYKSNACKKHQRIDPRKSTFONLGEPLSIHYGTSMQSLIDVTVSNIVDIQT
VGLSTQEPGVFTYAFEDGILMAVPSLSEYSI PVFDMNMNHLVAODLFSYMDRI
GQESMLTGLAINPSYVTGSLWVPTVQPTWQPTVDSVTSGVVAACEGCGAILDTG
TSLVGPSSDILNIQOALIGATONOGYGFIDCNLSYMPVVEINGKVPPLPSAYT
SODQGFCTSGFQSEHNSQKWLIDVFIREVYSVFDRAANLVGLAKAI"
misc_difference 1291
/gene="preprochymosin"
/note="POLY A oligo C"

BASE COUNT 309 a 384 c 334 g 263 t 1 others
ORIGIN

Query Match 76.2%; Score 894; DB 6; Length 1291;
Best Local Similarity 87.5%; Pred. No. 1.3e-263;
Matches 978; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 56 TCGTTCGTGTTACTCAACGCTGTGAGATCAACCGCATTCCTCTCAAAAGTAAGTCTC 115
DB 54 TCTTCGCTCTCTCCCAAGGCGCTGAGATCACAGAGTCCCTCTGTACAAAGGCAAGTCTC 113
QY 116 TCCGTAAGGCGCTGAAGGAATGACTTCTAGAAAGTCTTTCAGAAAACAAGATATG 175
DB 114 TGAGAAAGGCGCTGAAGGAATGACTTCTGAGAGACTTCTGCAAAAACAAGATATG 173
QY 176 GCATCAGCAGCAAGTACTCCGGCTTCGCTGAAGTTGCTAGCGTCCACTTAACAATACC 235
DB 174 GCATCAGCAGCAAGTACTCCGGCTTCGCTGAAGTTGCTAGCGTCCACTTAACAATACC 233
QY 236 TTGATGTCAATCTTTGGGAAGATCAACCTCGGAACCCCGGCTCAAGAGTTCAACGCTTC 295
DB 234 TAGATGTCAATCTTTGGGAAGATCAACCTCGGAACCCCGGCTCAAGAGTTCAACGCTTC 293

```

```

QY 296 TCTTGATATCTGCTTCTCTGACTTCTGGTTCCTCTATCTACTGCAAGAGATGCTT 355
DB 294 TCTTGATATCTGCTTCTCTGACTTCTGGTTCCTCTATCTACTGCAAGAGATGCTT 353
QY 356 GCAAGAACCAACAAAGATTCGATCCGAGAAAGTCGCACTTCCGAAGATTAGGCAAC 415
DB 354 GCAAGAACCAACCAAGATTCGATCCGAGAAAGTCGCACTTCCGAAGATTAGGCAAC 413
QY 416 CTTGTCATATACATACGATACAGGATGATGATGATGATGATGATGATGATGATGAT 475
DB 414 CTTGTCATATACATACGATACAGGATGATGATGATGATGATGATGATGATGATGAT 473
QY 476 CTGCTCCAAATTTGTCATTTCAACAGACAGTGAAGATTAGCAACCCCAAGAACGGTG 535
DB 474 CTGCTCCAAATTTGTCATTTCAACAGACAGTGAAGATTAGCAACCCCAAGAACGGTG 533
QY 536 ATGCTTTCACCTATGAGAAATTCAGATGATGATGATGATGATGATGATGATGATGAT 595
DB 534 ACCTTTCACCTATGAGAAATTCAGATGATGATGATGATGATGATGATGATGATGAT 593
QY 596 CAGAGTACTGATACCTGTTGTTGACACATGATGATGATGATGATGATGATGATGATGAT 655
DB 594 CAGAGTACTGATACCTGTTGTTGACACATGATGATGATGATGATGATGATGATGATGAT 653
QY 656 TGTTCGCTTTACATGACAGGAATGCGCAGAGAGCATGCTACGCTTGGAGCTATTG 715
DB 654 TGTTCGCTTTACATGACAGGAATGCGCAGAGAGCATGCTACGCTTGGAGGCGCATCA 713
QY 716 ATTCATCTACTACACAGGATCTTCACTGGGTTCACTGCTGACGACAGTACGAGT 775
DB 714 ACCGTCCTACTACACAGGATCTTCACTGGGTTCACTGCTGACGACAGTACGAGT 773
QY 776 AATTCATGTCGACAGTACCATCAGCGGTGTTGATGATGATGATGATGATGATGATGAT 835
DB 774 AGTTCACTGTCGACAGTACCATCAGCGGTGTTGATGATGATGATGATGATGATGATGAT 833
QY 836 AAGCTATCTTGTGATACCGGTAGTCCAGCTGTGCGAAGCTTACAGCAGCAATTTCAACA 895
DB 834 AGGCTATCTTGTGATACCGGTAGTCCAGCTGTGCGAAGCTTACAGCAGCAATTTCAACA 893
QY 896 TTGAGAACGATATGAGGACACACAGACAGATGATGATGATGATGATGATGATGATGAT 955
DB 894 TCCAGAGGCAATGAGGACACACAGACAGATGATGATGATGATGATGATGATGATGAT 953
QY 956 ACCTTACTACATGCTTACAGTGTCTTTGATCAACGCGCAAGATGATCCCACTGACCC 1015
DB 954 ACCTGAGCTACATGCTTACAGTGTCTTTGATCAATGCAAAAGTACCACTGACCC 1013
QY 1016 CTTCCGCTATACGACGATCAAGGCTTGTGACCAAGTGAATTCAGAGTGAAC 1075
DB 1014 CTTCCGCTATACGACGATCAAGGCTTGTGACCAAGTGAATTCAGAGTGAAC 1073
QY 1076 ATTCGCAAGAAATGATCTGGAGATGTGTTCAATGCTGATCTACAGGCTTTGACA 1135
DB 1074 ATTCGCAAGAAATGATCTGGAGATGTGTTCAATGCTGATCTACAGGCTTTGACA 1133
QY 1136 GGGCCACAACCTCTGTTGGCTGCTAAGCAATCTGA 1173
DB 1134 GGGCCACAACCTCTGTTGGCTGCTAAGCAATCTGA 1171

RESULT 12
E00144 1289 bp RNA linear PAT 29-SEP-1997
LOCUS E00144
DEFINITION cdna encoding bovine calf chymosin.
ACCESSION E00144
VERSION E00144.1 GI:2168443
KEYWORDS JP 1984021392-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1289)

```


AUTHORS Chiyaarusu, E.B.
TITLE CATTLE KIMOCIN
JOURNAL Patent: JP 1984021392-A 1 03-FEB-1984;

FEATURES	Feature	Population
Location/Qualifiers	Location/Qualifiers	Population

Query Match	76.1%	Score 893	DB 6	Length 1289
Best Local Similarity	87.4%	Pred No 2.6e-263		
Matches 977	Conservative 0	Mismatches 141	Indels 0	Gaps 0
QY	56	TCGTGCTGTTACTACGCGCTGCTGAGATCACCGCATTCCTCTTACAAAGGTAAAGTCTC	115	
Db	23	TCCTTCGCTCTCTCCACGAGGCGCTGAGATCACAGAGATCCCTCTGTACAAAGGCAAGTCTC	82	
QY	116	TCGCGAAGGCGCTGAAAGAACATGGAACCTTATGAAGACCTTGTGAGAAACAAACGATAG	175	
Db	83	TCGAGAAAGCGCTGAAGAGACATGGAGCTTCTGGAGAGACTTCTGCAAGAAACGACGTATG	142	
QY	176	GCATCAGCAGCAAGTACTCCGGCTTCGGTGAAGTTGTAACGTGCACTTACCAACTAC	235	
Db	143	GCATCAGCAGCAAGTACTCCGGCTTCGGGAGAGTGCGCAGCGTGCCTCCGACCAACTAC	202	
QY	236	TTGATATGCACTACTTTGGGAAGATCTACCTCGAAACCCCGGCTCAAGATTCACCGTTC	295	
Db	203	TAGATATGCACTACTTTGGGAAGATCTACCTCGGAACCCCGGCTCAAGATTCACCGTTC	262	
QY	296	TCCTTGAATACGTGCTCCTCTGACTTCTGGGTCCTCTATCTATGCAAGACAAATGCT	355	
Db	263	TGTTTGAACATGGGCTCCTCTGACTTCTGGGTAACCTCAATCTATGCAAGACAAATGCT	322	
QY	356	GCAAGAACCAACCAAGATTCGATCCGAGAAAGTGGTCCACCTTCCAGAACTTAGGCAAC	415	
Db	323	GCAAAACCAACCAAGGCTTCGACCCGAGAAATCTCTCACCTTCCAGAACTTGGGCAAGC	382	
QY	416	CCTTGCTATACACTACGCTACAGGTAGATGATCAAGAAATCTTAGGCTATGATACCGTCA	475	
Db	383	CCCTGTCTATCCACTACCGGACAGGCAAGATCAAGGATCTCTGGCTATGACACCGTCA	442	
QY	476	CTGTCTCCAACTTGTGACATTTCAACACACATGAGACTTATGACCCCAAGAACCAAGGTG	535	
Db	443	CTGTCTCCAACTTGTGACATTTCAACACACATGAGGCTTAGACACCAAGAGCCCGGGG	502	
QY	536	ATGCTTACCACTATGCAATTCATGAGCATCTTGATGTGACATACCATGCTGCGGT	595	
Db	503	ACGTCTTACCACTATGCAATTCATGAGGATCTCTGGGATGCGCTTACCCCTGCTGCGCT	562	
QY	596	CAGAGTACTCGATACCTGTGTTGACAAACATATGACCGACACCTTAGTACTCAAGACT	655	

Db	563	CAGAGTACTGTGATTAACCCGTGTGTTGACAACATGATGAACAGGACCTGTGTGGCCAGAAC	622			
Qy	656	TGTTCTGTGGTTTACATGAGACAGAAATGGCCAGGAGACATCTTCACGCTTGGAGCTAAT	715			
Db	623	TGTTCTCGATTTACATGAGACAGAAATGGCCAGGAGACATGCTTCACGCTGGGGGCCAT	682			
Qy	716	ATCATCTCTACTACACAGGATCTCTTTCATCTGGGTTTCCAGTCACTGTGCACAGTACTGGC	775			
Db	683	ACCGTCTCTACTACACAGGATCTCTTTCATCTGGGTTTCCAGTCACTGTGCACAGTACTGGC	742			
Qy	776	AATTCACCTGTGAGACAGTGTACACCATTCAGCGGCTGTGTGTTGCACTGTGAAGGTTGATGTC	835			
Db	743	AGTTCACTGTGTGAGACAGTGTACACCATTCAGCGGCTGTGTGTTGCACTGTGAAGGTTGTC	802			
Qy	836	AAGCTATCTTGTGATTCACCGTACGTCTTTCAGGCTGTGTGACCTTCAGGACGACATTTCTCA	895			
Db	803	AGGCGATCTGTGAGACAGGGGACCTCCAGAGTGTGTGGGCCACAGAGGACATCTCTCA	862			
Qy	896	TTTCAGCAAGCTATTGAGCCACACAGAACGATACGGTGAAGTTGACATGATTTGCGACA	955			
Db	863	TTCAGCAGGGCCATTGGAGCCACACAGAACGATACGGTGAAGTTGACATGATTTGCGACA	922			
Qy	956	ACCTTACTCTACATGCGCCACAGTTGTCTTTGATTCACACGGCAAGATGTATCCACTGACCC	1015			
Db	923	ACCTGACTCTACATGCGCCACAGTTGTGTCTTTGATTCACATGCGCAAAATGTATCCACTGACCC	982			
Qy	1016	CTTCCGCTTATTCACGACGACGATTCAGAGGTTCTGCAACAGTGTATTCAGAGTGAAC	1075			
Db	983	CTTCCGCTTATTCACGACGACGATTCAGAGGCTTCTGTACATGTGCTTCCAGAGGAAATC	1042			
Qy	1076	ATTCCCAAGAAATGAGATCTTGGAGATGTGTCAATTCGTGAGTACATACAGCGTCTTTGACA	1135			
Db	1043	ATTCCCAAAATGAGATCTTGGAGGATGTGTTCATTCAGAGTATTCACAGCGTCTTTGACA	1102			
Qy	1136	GGGCCAACAACTCTGTGGGCTAGCTTAACCAATCTGA	1173			
Db	1103	GGGCCAACAACTCTGTGGGCTAGGCCAAGCCATCTGA	1140			
RESULT 13						
LOCUS	108097	108097	2726 bp DNA linear PAT 02-DEC-1994			
DEFINITION	Sequence 5 from Patent EP 0301669.					
ACCESSION	108097					
VERSION	108097.1	GI:589192				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 2726)					
AUTHORS	van den Berg,J.A.D. and Brake,A.J.D.					
TITLE	DNA constructs containing a Kluveromyces alpha-factor leader					
JOURNAL	sequence for directing secretion of heterologous derivatives					
FEATURES	Patent: EP 0301669-A1 5 01-FEB-1989;					
source	1..2726					
BASE COUNT	751 a 654 c 572 g 749 t					
ORIGIN	/organism="unknown"					
Query Match 76.1%; Score 893; DB 6; Length 2726;						
Best Local Similarity 87.5%; Pred. No. 2,7e-263;						
Matches 977; Conservative 0; Mismatches 140; Indels 0; Gaps 0;						
Qy	57	CGTTGCTGTACTACAGCTGTGTGATCACCCGATTTCTCTTACAAAGTAACTTCT	116			
Db	669	CGATGCTTCCATCATATGCTGTGATCACAGATCCCTGTATCAAAAGCAAGTCTCT	728			
Qy	117	CCGTAAGCGCTGAAGAACATGGAATCTTCTAGAAAGATTTCTTCAGAAAACAAGATATG	176			
Db	729	GAGGAAGCGCTGAAGAGCATGGCTTCTGAGAGATTTCTCTGAGAAAACAAGCATATG	788			
Qy	177	CATAGCAGACGACTCCGGCTTCGGTGAAGTTGTCTAGCGTCCACTTACCAATCACT	236			

RESULT 13									
108097	LOCUS	108097	2726 bp	DNA	linear	PAT 02-DEC-1994			
108097	DEFINITION	Sequence 5 from Patent EP 0301669.							
108097	ACCESSION	108097							
108097.1	VERSION	GI:589192							
	KEYWORDS	.							
SOURCE		Unknown.							
ORGANISM		Unknown.							
REFERENCE		Unclassified.							
AUTHORS		1 (bases 1 to 2726)							
TITLE		van den Berg,J.A.D. and Brake,A.J.D.							
JOURNAL		DNA constructs containing a Kluweromyces alpha-factor leader							
FEATURES		sequence for directing secretion of heterologous derivatives							
source		Patent: EP 0301669-A1 5 01-FEB-1989;							
		Location/Qualifiers							
		1..2726							
BASE COUNT		751 a	654 c	572 g	749 t				
ORIGIN		/organism="unknown"							
Query Match		76.1%;	Score 893;	DB 6;	Length 2726;				
Best Local Similarity		87.5%;	Pred. No. 2.7e-263;						
Matches 977;		Conservative 0;	Mismatches 140;	Indels 0;	Gaps 0;				
QY	57	CGTGGCTGTACTACGACGCTGTGAGATCCACCCGATTCCTCTCTACCAAGTAACTTCT	116						
Db	669	CGATGCTTCCATCAATVAGCTGTGATCAACAGATCCCTCTGTACAAAGCAAGTCTCT	728						
QY	117	CCGTAAGCGCCTGAAGAACATGACATCTTCTAGAAAGATCTTGTGCAGAAACAAAGATATG	176						
Db	729	GAGAAAGCCGCTGAAGAGCATGGGCTTCTGAGAGATTTCTGACAAAACAGCAGTATGG	788						
QY	177	CATCAGCAGCAAGTACTCCGAGCTTCGGTGAAGTTGCTAGCGTGCACCTTACCAACTACCT	236						

D	789	CATCAGCAGAGTACTCCGGCTTCGGGGAGGTGGCCAGGTCCTCCCTGACCACTACT	848
Q	237	TGATAGTCAATCTTTGGGAAAGTCTAAGTCAAGTCCGCTCAAGAGTTCAACGTTCT	296
D	849	GGACAGTCACTTTGGGAAAGTCACTCGGGAACCCGCCAGAGATTCAACCGTCT	908
Q	297	CTTTGATACAGTCTTCCTGAGCTTTGGGTTCCCTCTATCTACTGCAAGAGCAATGCTG	356
D	909	GTTTGACACTGGCTCTGAGCTTCGGGTACCCCTATCTACTGCAAGAGCAATGCTG	968
Q	357	CAAGAACAACCAAGATTGATCCGAGAAAGTGTCCACCTTCAGAACTTAAAGCAAAAC	416
D	969	CAAAAACCAACAGCGCTTCGACCCGAGAAAGTGTCCACCTTCAGAACTTCGGCAAGCC	1028
Q	417	CTTGTCTATACACTAGCGGTAACAGTATGATGCAAGAACTTAAAGCTATGATACCGTCA	476
D	1029	CTGTCTATCACTAGCGGTAACAGGATGATGCAAGGATCTGCGGTATGACACCGTCA	1088
Q	477	TGTCTCAACATTTGGAGCACTTCAACAGAGTATGCACTTAAAGCAAGCAAGAGTGA	536
D	1089	TGTCTCAACATTTGGAGCACTTCAACAGAGTATGCACTTAAAGCAAGCAAGAGTGA	1148
Q	537	TGTCTCAACATTTGGAGCACTTCAACAGAGTATGCACTTAAAGCAAGCAAGAGTGA	596
D	1149	CGTCTCAACATTTGGAGCACTTCAACAGAGTATGCACTTAAAGCAAGCAAGAGTGA	1208
Q	597	AGAGTACTGATACCTGTGTTTGAACAATGATGAAACCAACCTAGTATGCTCAAGATT	656
D	1209	AGAGTACTGATACCTGTGTTTGAACAATGATGAAACCAACCTAGTATGCTCAAGATT	1268
Q	657	GTTCTGGTTTAAATGAGCAAGAAATGGCCAGAGAGATGCTCAACGTTTGGAGTATGA	716
D	1269	GTTCTGGTTTAAATGAGCAAGAAATGGCCAGAGAGATGCTCAACGTTTGGAGTATGA	1328
Q	717	TCCATCTTATACAGAGATCTCTTCACTGAGTTTCAAGTCACTGTGAGCAAGTATGCA	776
D	1329	CCGCTCTTATACAGAGATCTCTTCACTGAGTTTCAAGTCACTGTGAGCAAGTATGCA	1388
Q	777	ATTCACTGTGAGCAAGTATGCACTGAGTTTCAAGTCACTGTGAGCAAGTATGCA	836
D	1389	GTTCACTGTGAGCAAGTATGCACTGAGTTTCAAGTCACTGTGAGCAAGTATGCA	1448
Q	837	AGCTATCTTGGATACCGGTAAGTCAAGTCTGTGAGCACTTAAAGCAATCTTCAAT	896
D	1449	GAGCACTCTGGAGCAAGGCACTTCAAGTCTGTGAGCACTTAAAGCAATCTTCAAT	1508
Q	897	TGACAGAGTATTTGAGGCAACAAGACAGTACGATGAGTTTGAAGTTTGGCAAA	956
D	1509	CGAGCAAGCAATTTGAGGCAACAAGACAGTACGATGAGTTTGAAGTTTGGCAAA	1568
Q	957	CCTTACGATCATGCTCAAGTTGTCTTGAAGTCAAGGCAAGTATGCACTGACCC	1016
D	1569	CTGAGATCATGCTCAAGTTGTCTTGAAGTCAAGGCAAGTATGCACTGACCC	1628
Q	1017	CTCCGCTTATACAGGCAAGATGATGAGTCTGCAAGTATGCAAGTATGCAAGTGA	1076
D	1629	CTCCGCTTATACAGGCAAGATGATGAGTCTGCAAGTATGCAAGTATGCAAGTGA	1688
Q	1077	TTCCCAAGAAATGATCTTTGGAGATGATGATCTGATGAGTATGCAAGGCTTTGACAG	1136
D	1689	TTCCCAAGAAATGATCTTTGGAGATGATGATCTGATGAGTATGCAAGGCTTTGACAG	1748
Q	1137	GAGCAACCACTGTGAGGCTAGTAAAGCAATCTGA	1173
D	1749	GAGCAACCACTGTGAGGCTAGTAAAGCAATCTGA	1785

RESULT 14
E00108
LOCUS E00108 1311 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA coding of prepro chymosin.
ACCESSION E00108

VERSION	E00108.1	GI:2168409
KEYWORDS	JP 1983109499-A/1.	
SOURCE	unidentified	
ORGANISM	unclassified	
REFERENCE	1 (bases 1 to 1311)	
AUTHORS	Yan, M., Kornueerisu, S.B., Adorianusu, M.R. and Rutsupo, E.	
TITLE	RECOMBINED DNA	
JOURNAL	Patent: JP 1983109499-A 1 29-JUN-1983;	
COMMENT	UNILEVER NV	
OS	calif	
PN	JP 1983109499-A/1	
PD	29-JUN-1983	
PR	14-OCT-1982 JP 1982180549	
PR	14-OCT-1981 GB 81 8131004	
PI	YAN MAATO, KORUNEEERISU SEODORUSU BEERUTSUPUSU, PI	
ADRIANUSU MARINUSU REDEBERU, RUTSUPO EDENSU PC		
CC	C07H21/04, C12N1/20, C12N15/00//C12P21/00;	
CC	strandedness: Double;	
CC	topology: Linear;	
CC	hypothetical: No;	
CC	anti-sense: No;	
CC	*source: tissue_type=stomach;	
PH	Key	Location/Qualifiers
PH		
FT	CDS	24..1169
FT		/product='prepro chymosin'
FT	sig_peptide	24..71
FT	mat_peptide	72..1166
FT		/product='prochymosin'
FT	mat_peptide	153..1166
FT		/product='pseudochymosin'
FT	mat_peptide	198..1166
FT		/product='chymosin'
FEATURES	source	1..1311
	Location/Qualifiers	
BASE COUNT	309 a 397 c 338 g 267 t	
ORIGIN		
Query Match	76.1%;	Score 892.4; DB 6; Length 1311;
Best Local Similarity	87.4%;	Pred. No. 3.9e-263;
Matches	977; Conservative	0; Mismatches 141; Indels 0; Gaps 0;
Q	56	TCGTGCTGTTATCAACGCTGCTGAGATCAACCGCATTCCTCTCAAAAGTAAGTCTC 115
D	52	TCTTCGCTCTCTCCCAAGCGCTGAGATCAACGATTCCTCTGTACAAAGCAAGTCTC 111
Q	116	TCCGTAAGCGCTGAAGAAATGATGACTTCTAGAAAGTCTTGGCAAAACACAGTATG 175
D	112	TGAGAAAGCGCTGAAGAGCATGAGCTTCTGAGAGACTTCTGCAAAACACAGTATG 171
Q	176	GCATGACGACAGTACTCCGCTTGGTGAAGTCTGAGCGTCCACTTAACTTAACTTACC 225
D	172	GCATGACGACAGTACTCCGCTTGGGAGGTGCGACCGCTGACCAACTTACC 221
Q	236	TGATAGTCAATCTTTGGGAAAGTCTACTCGGAACCCGCTCAAGAGTTCAACGTTG 295
D	232	TGATAGTCAATCTTTGGGAAAGTCTACTCGGAGACCCGCTCAAGAGTTCAACGTTG 291
Q	296	TCTTTGATATGAGTCTCTGACTTCTGAGTTTCCCTCTATCTAAGTCAAGAGTATGCT 355
D	292	TGTTTGAAGTCTCTCTGACTTCTGAGTTTCCCTCTATCTAAGTCAAGAGTATGCT 351
Q	356	GCAAGAACCAACCAAGTTGATCCGAGAAAGTCTTCACTTCCAGAACTTAAAGCAAAAC 415
D	352	GCAAGAACCAACCAAGTTGATCCGAGAAAGTCTTCACTTCCAGAACTTAAAGCAAAAC 411
Q	416	CTTGTCTATACACTAGCGGTAACAGTATGCAAGAACTTAAAGCTATGATACGTTCA 475
D	412	CCCTGTATCACTAGCGGTAACAGGATGCAAGGATCTGAGGCTATGACACCGTCA 471


```
QY      656 TGTTCGTTTACATGACAGGAATGGCCAGGAGCATGCTCAGCCTTGAGCTATTG 715
      |||
Db      833 TGTTCGTTTACATGACAGGAATGGCCAGGAGCATGCTCAGCCTTGAGCTATTG 892
      |||
QY      716 ATTCATCTACTACACAGGATCTTTCATGGGTTCCAGTCACTGTGCAGCAGTACTGGC 775
      |||
Db      893 ACCGTCCTACTACACAGGATCTTTCATGGGTTCCAGTCACTGTGCAGCAGTACTGGC 952
      |||
QY      776 AATTCACTGTGACAGTGTACCATCAAGCCGTGTGTTGTTGCATGTGAAGTGTGATGTC 835
      |||
Db      953 AGTTCACTGTGACAGTGTACCATCAAGCCGTGTGTTGTTGCATGTGAAGTGTGATGTC 1012
      |||
QY      836 AAGCTATCTTGGATTACCGGTACGATCCAGCTGGTGGACCTAGCAGGACATTCCTCAACA 895
      |||
Db      1013 AAGCCATCTTGGATTACCGGTACGATCCAGCTGGTGGACCTAGCAGGACATTCCTCAACA 1072
      |||
QY      896 TTCAGCAAGCTATTGGAGCCACAAGAACCAAGTACGGTGAATTGACATAGATTGCGACA 955
      |||
Db      1073 TTCAGCAAGCTATTGGAGCCACAAGAACCAAGTACGGTGAATTGACATAGATTGCGACA 1132
      |||
QY      956 ACCTTAGCTACATGCTTACAGTTGTCTTTGAGATCAACGGCAAGATGTACCACTGACCC 1015
      |||
Db      1133 ACCTTAGCTACATGCTTACAGTTGTGTGAGATCAATGGCAAAATGTACCCAGTGACCC 1192
      |||
QY      1016 CCTCCGCTATACAGCAGGATCAAGGTTCTGCACCAAGTGAATCCAGAGTGAAGACC 1075
      |||
Db      1193 CCTCCGCTATACAGCAGGATCAAGGTTCTGTACCAAGTGAATCCAGAGTGAAGATC 1252
      |||
QY      1076 ATTCAGAAATGATCTTGGAGATGTGTTCATTGCTGAGTACTACAGCGCTTTGACA 1135
      |||
Db      1253 ATTCAGAAATGATCTTGGAGATGTGTTCATTGAGAGTATTACAGCGCTTTGACA 1312
      |||
QY      1136 GGGCCAAACAACCTGCTGGGCTAGCTAAAGCAATCTGA 1173
      |||
Db      1313 GGGCCAAACAACCTGCTGGGCTAGCCAAAGCAATCTGA 1350
      |||
```

Search completed: February 5, 2004, 04:30:01
Job time : 4602 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2004, 01:27:00 ; Search time 390 Seconds
(without alignments)
8119.079 Million cell updates/sec

Title: US-09-643-755B-1

Perfect score: 1173
Sequence: 1 atgacatcctaagctctt.....ggctagctaaacatctga 1173

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1173	100.0	1173	22 AAS00569
2	1173	100.0	3957	22 AAS00570
3	898.8	76.6	1291	10 AAN91157
4	897.2	76.5	1175	5 AAN40295
5	897.2	76.5	1240	16 AAT03006
6	897.2	76.5	1460	5 AAN40180
7	895.6	76.4	1278	5 AAN40055
8	895.6	76.4	1460	3 AAN20043

9	894.8	76.3	2733	20 AAZ06463
10	894.8	76.3	2733	24 ABS53073
11	894	76.2	1210	12 AAO14051
12	894	76.2	1290	4 AAN30209
13	893	76.1	2727	10 AAN91188
14	890.8	75.9	1314	4 AAN30049
15	889.8	75.9	1289	4 AAN30022
16	886.8	75.6	1098	11 AAO04683
17	885.2	75.5	1098	4 AAN30063
18	872.8	74.4	2982	10 AAN91185
19	866.6	73.9	1175	13 AAO20949
20	716.8	61.1	1146	24 ABS53735
21	704.6	60.1	1143	5 AAN40214
22	686.8	58.6	1104	9 AAN80001
23	534	45.5	1140	24 AAS97159
24	534	45.5	1238	24 ABS58369
25	382.4	32.6	637	14 AAO49459
26	359.2	30.6	1393	22 AAS7421
27	359.2	30.6	1409	22 AAS44608
28	315.4	26.9	1017	23 AAS79579
29	300.8	25.6	1210	24 AAD37045
30	290	24.7	1360	20 AAZ20193
31	253	21.6	2158	25 ACC50980
32	253	21.6	2158	25 ABX76363
33	253	21.6	2201	24 ABZ35346
34	248.4	21.2	1357	22 AAS7422
35	248.4	21.2	1365	22 AAD09486
36	248.4	21.2	1366	24 ABK64746
37	248.4	21.2	1366	24 ABL69575
38	248.4	21.2	1418	20 AAZ71523
39	247.6	21.1	1340	20 AAZ20179
40	235.2	20.1	1285	20 AAZ20181
41	233.6	19.9	1258	20 AAZ20162
42	231.6	19.7	1130	20 AAZ20182
43	230.4	19.6	1173	20 AAZ20183
44	226.8	19.3	1173	24 AAD30576
45	226.8	19.3	1173	24 AAS97161

ALIGNMENTS

RESULT 1			
AA00569	standard: DNA, 1173 BP.		
AC	AA00569;		
XX			
XX	14-MAY-2001 (first entry)		
DT			
XX			
DE	Bovine pre-pro-chymosin DNA sequence.		
XX			
KW	Chymosin; transcription regulator; terminator sequence; soybean; corn;		
KW	pre-pro-chymosin; rape seed; sunflower; cotton; tobacco; alfalfa; wheat;		
KW	barley; oats; soybean; Arabidopsis thaliana; potato; flax; linseed; rice;		
KW	safflower; oil palm; ground nut; Brazil nut; coconut; castor; coriander;		
KW	squash; jojoba; ds.		
OS	Bos sp.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	1..1173	
FT		/tag= a	"Bovine chymosin"
FT		/product=	
FT	sig_peptide	1..78	
FT		/tag= b	
FT	misc_feature	79..201	
FT		/tag= c	
FT		/note= "Pro sequence"	
FT	mat_peptide	202..1170	
FT		/tag= d	
FT		/product=	"Mature bovine chymosin"
XX			

2.7 Kbp HindIII fr
DNA encoding oleos
Remain gene. Synt
Sequence of prepro
BamHI/SalI insert
cDNA sequence corr
Sequence of veal c
Sequence encoding
Sequence of proren
BamHI insert from
Prochymosin (prore
Aspergillus gene e
Gene encoding (pre
Optimised prochymo
Human aspartyl pro
protein modificati
Prochymosin gene 5
Human stomach cell
Human full-length
DNA encoding novel
Mouse pepsinogen-f
Cat pregnancy asso
Human bladder canc
Lung cancer-associ
Human gene express
Human stomach cell
Human gastricsin c
Human benign prost
Prostate cancer re
Human ovarian tumo
Bovine pregnancy a
Bovine pregnancy a
Bovine pregnancy a
Bovine pregnancy a
Human proteinase, PR
Human aspartyl pro

XX	MOZ00114571-AL.
XX	
PB	01-MAR-2001.
XX	
PF	23-AUG-2000; 2000WO-CA00975.
XX	
PR	23-AUG-1999; 99US-0378696.
XX	
PA	(SEMB-) SEMBIOSYS GENETICS INC.
XX	
PI	Van Rooijen G, Keon RG, Boeche J, Shen Y;
XX	
DR	WPI: 2001-226521/23.
XX	
PT	P-RSDB; AAU00536.
XX	
FT	Producing chymosin in seeds of plants such as rice, flax, rape seed, by transforming plant cell with a nucleic acid encoding chymosin operably linked to transcription regulator and terminator sequences -
XX	
PS	Claim 9; Fig 1; 56pp; English.
XX	
CC	The sequence represents a DNA which encodes a bovine chymosin polypeptide. Chymosin can be produced in a plant seed through introduction of a chimeric nucleic acid molecule, comprising a nucleic acid sequence encoding a chymosin polypeptide operatively linked to a transcription regulator and terminator sequences, into a plant cell. The sequences are useful for producing plant seeds, in particular seeds of soybean, rape seed, sunflower, cotton, corn, tobacco, alfalfa, wheat, barley, oats, sorghum, Arabidopsis thaliana, potato, flax/lilseed, safflower, oil palm, groundnut, Brazil nut, coconut, castor, coriander, squash, jojoba and rice.
XX	
SQ	Sequence 1173 BP, 299 A; 308 C; 262 G; 304 T; 0 other;
XX	
Query Match	100.0%; Score 1173; DB 22; Length 1173;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1173; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	1 ATGAACCTTCCTTAAGTGTCTTTCCCTTTCTAAGGCTTTCTTTGGTTGGATACATTCTCGTT 60
Db	1 ATGAACCTTCCTTAAGTGTCTTTCCCTTTCTAAGGCTTTCTTTGGTTGGATACATTCTCGTT 60
QY	61 GCTGTACTCAGCGTCGTGAATCACGCCGAACTTCTCTATAAAGTAAAGTCTCTCGGT 120
Db	61 GCTGTACTCAGCGTCGTGAATCACGCCGAACTTCTCTATAAAGTAAAGTCTCTCGGT 120
QY	121 AAGGCGCTGAAGAATCATGGAATTTCTGAAGAATCTTGSCAGAAAACAAGTAGGCATC 180
Db	121 AAGGCGCTGAAGAATCATGGAATTTCTGAAGAATCTTGSCAGAAAACAAGTAGGCATC 180
QY	181 AGACGCAAGTACTCCGCGCTTCGGGTGAAGTTGCTAGCGTCCACTTAACCACTACCTTGAT 240
Db	181 AGACGCAAGTACTCCGCGCTTCGGGTGAAGTTGCTAGCGTCCACTTAACCACTACCTTGAT 240
QY	241 AGTCAATACTTTGGGAAGATCTAAGTCCGGAACCCGCGCTCAAGAAGTTCAACCGTTCTTT 300
Db	241 AGTCAATACTTTGGGAAGATCTAAGTCCGGAACCCGCGCTCAAGAAGTTCAACCGTTCTTT 300
QY	241 AGTCAATACTTTGGGAAGATCTAAGTCCGGAACCCGCGCTCAAGAAGTTCAACCGTTCTTT 300
Db	241 AGTCAATACTTTGGGAAGATCTAAGTCCGGAACCCGCGCTCAAGAAGTTCAACCGTTCTTT 300
QY	301 GATATCTGGTCTCTGACTCTTGGGTTCCCTCTATCTACTGCAAGCAATGCTGGCAAG 360
Db	301 GATATCTGGTCTCTGACTCTTGGGTTCCCTCTATCTACTGCAAGCAATGCTGGCAAG 360
QY	361 AACCAACCAAAGATGTGATCCGAAAGTGTGCACCTTCAGAACTTAGCAAAACCTTGG 420
Db	361 AACCAACCAAAGATGTGATCCGAAAGTGTGCACCTTCAGAACTTAGCAAAACCTTGG 420
QY	421 TCATATACACTACGGTACAGGTAGCATGCAAGAAATCTTAGGCTATGATACCGTCACTGTC 480
Db	421 TCATATACACTACGGTACAGGTAGCATGCAAGAAATCTTAGGCTATGATACCGTCACTGTC 480
QY	481 TTCACATTTGGAGCAATTCAACAGACAGTAGAATTAGACCACCAAGAACCAAGTAGATGTC 540
Db	481 TTCACATTTGGAGCAATTCAACAGACAGTAGAATTAGACCACCAAGAACCAAGTAGATGTC 540

OY	541	TTCACTCATGACGAATTCGATATGGCATCTCTTGGATATGGCATATACCATATGGCTTCGGCTCAGAG	600
Db	541	TTCACTCATGACGAATTCGATATGGCATCTCTTGGATATGGCATATACCATATGGCTTCGGCTCAGAG	600
OY	601	TACTCGATACCTGTGTTTGAACAACATGATGAACCGACACCTAGTAGCTCAAGAATTGTC	660
Db	601	TACTCGATACCTGTGTTTGAACAACATGATGAACCGACACCTAGTAGCTCAAGAATTGTC	660
OY	661	TCGGTTTACATGACAGAGAAATGGCCAGAGAGCATGCTCAAGCTTGGAGCTATTGATCCA	720
Db	661	TCGGTTTACATGACAGAGAAATGGCCAGAGAGCATGCTCAAGCTTGGAGCTATTGATCCA	720
OY	721	TTCCTACTACACAGGATCTCTTTCACCTGGGTTCCAGTCACTGTGACAGAGTACCTGGCAATTC	780
Db	721	TTCCTACTACACAGGATCTCTTTCACCTGGGTTCCAGTCACTGTGACAGAGTACCTGGCAATTC	780
OY	781	ACTGTGACAGATGTCAACCATCAGCGAGTGTGGATTGTTCATATGTGAAGGTGATGTCAAGCT	840
Db	781	ACTGTGACAGATGTCAACCATCAGCGAGTGTGGATTGTTCATATGTGAAGGTGATGTCAAGCT	840
OY	841	ATCTTGGATACCGGTACGTCCAGCTGGTGGACCTTAGCACGCCGATTTCTACATTTCAAG	900
Db	841	ATCTTGGATACCGGTACGTCCAGCTGGTGGACCTTAGCACGCCGATTTCTACATTTCAAG	900
OY	901	CAAGCTATTGGAGGCCACACAGAAACAGTACGGTGAATTTGACATAGATTTGGACAAACCTT	960
Db	901	CAAGCTATTGGAGGCCACACAGAAACAGTACGGTGAATTTGACATAGATTTGGACAAACCTT	960
OY	961	AGTACATATGCCCTACAGTGTCTTTGAGATCAACCGGCAAGATGTACCCATGACCCCTCC	1020
Db	961	AGTACATATGCCCTACAGTGTCTTTGAGATCAACCGGCAAGATGTACCCATGACCCCTCC	1020
OY	1021	GCTATATACCAAGCCAGGATCAAGGGTTCTGCACACGTGATTTCCAGAGTGAAGACATTTCC	1080
Db	1021	GCTATATACCAAGCCAGGATCAAGGGTTCTGCACACGTGATTTCCAGAGTGAAGACATTTCC	1080
OY	1081	CAGAAATGGATCTTGGGAGAGATGTGTTCATTGCTGAGTACTACAGCGTCTTTTGA CAGGAGCC	1140
Db	1081	CAGAAATGGATCTTGGGAGAGATGTGTTCATTGCTGAGTACTACAGCGTCTTTTGA CAGGAGCC	1140
OY	1141	AACAACTCGTTGGGCTAGCTTAAGCAATCTGA	1173
Db	1141	AACAACTCGTTGGGCTAGCTTAAGCAATCTGA	1173
RESULT 2			
AAS00570			
XX	1D	AAS00570 standard; DNA; 3957 BP.	
XX	AC		
XX	AA	AAS00570;	
XX	DT	14-MAY-2001 (first entry)	
DE	XX	Bovine phaseolin promoter pre-pro-chymosin-phaseolin terminator.	
KM	XX	Chymosin; transcription regulator; terminator sequence; soybean; corn;	
KM	XX	pre-pro-chymosin; rape seed; sunflower; cotton; tobacco; alfalfa; wheat;	
KM	XX	barley; oats; sorghum; Arabidopsis thaliana; potato; flax; linseed; rice;	
KM	XX	safllower; oil palm; ground nut; Brazil nut; coconut; castor; coriander;	
KM	XX	squash; Jojoba; ds; phaseolin; promoter; terminator; mutant; French bean.	
OS	XX	Chimeric - Bos BP.	
OS	XX	Chimeric - Phaseolus vulgaris.	
XX	XX		
FH	Key	Location/Qualifiers	
FT	promoter	1..1553	
FT		/*tag= a	
FT		/note= "Phaseolin promoter"	
FT	CDS	1554..2726	
FT		/*tag= b	
FT		/product= "Bovine pre-pro-chymosin"	
FT	terminator	2727..3957	

[illegible]

OY		481	TTCGAACCTTGGGACATTTCAAAGCAAGTAGACTTGACGCCAAGAACCAGAGGATGTGC	540
Dd		2034	TTCAAACATTGGGACATTTCAAACAAGCAGTAGACATTAGAACCCAAAGAACCAAGGATGTGC	2093
OY		541	TTCACTCATGAGAAATTTCGATGGGATCTTGATATGGCATATACCATGCTCGCGTCAGAG	600
Dd		2094	TTCACTCATGAGAAATTTCGATGGGATCTTGATATGGCATATACCATGCTCGCGTCAGAG	2153
OY		601	TACTCGATACCTGTGTGTTGACAACATGATGAACCGAACCTAGTAGCTCAAAGCTTGTTC	660
Dd		2154	TACTCGATACCTGTGTGTTGACAACATGATGAACCGAACCTAGTAGCTCAAAGCTTGTTC	2213
OY		661	TCGGTTTTACATNGACAGAAATGGCCAGAGAGCATGTCOAGCTTGAGGCTATTTGATCCA	720
Dd		2214	TCGGTTTTACATNGACAGAAATGGCCAGAGAGCATGTCOAGCTTGAGGCTATTTGATCCA	2273
OY		721	TCCTACTACACAGAGATCTCTTCACTGGGTTCCAGTCACTGTGCAGCAGTACTGGCAATTC	780
Dd		2274	TCCTACTACACAGAGATCTCTTCACTGGGTTCCAGTCACTGTGCAGCAGTACTGGCAATTC	2333
OY		781	ACTGTGACAGGTGTCAACCATCAGCGGTGTGTTGTCATGTGAAGGTGATGTCAAGCT	840
Dd		2334	ACTGTGACAGGTGTCAACCATCAGCGGTGTGTTGTCATGTGAAGGTGATGTCAAGCT	2393
OY		841	ATCTTGGATACCGGTAAGTCCCAACCTGCTGGGAGCCTTAGACAGCGACATTTCCAACTTCAG	900
Dd		2394	ATCTTGGATACCGGTAAGTCCCAACCTGCTGGGAGCCTTAGACAGCGACATTTCCAACTTCAG	2453
OY		901	CAGCTATTATGGAGCACACAGAACCACTAGCGTAGTTTGACATAGATTGCGACAACTT	960
Dd		2454	CAGCTATTATGGAGCACACAGAACCACTAGCGTAGTTTGACATAGATTGCGACAACTT	2513
OY		961	AGTACATGECTACAGTTGTCTTGGATGACACGGCAAGATGTAACCACTGACCCCCTCC	1020
Dd		2514	AGTACATGECTACAGTTGTCTTGGATGACACGGCAAGATGTAACCACTGACCCCCTCC	2573
OY		1021	GCCATATCCAGCCAGAGATCAAGGGTCTGACACAGTGGATTCCAGAGTAGGAACCATTTCC	1080
Dd		2574	GCCATATCCAGCCAGAGATCAAGGGTCTGACACAGTGGATTCCAGAGTAGGAACCATTTCC	2633
OY		1081	CAGAAATGATCTTTGGAGAGATGTGTTCATTCTGTGAGTACTACAGCGTCTTTGACAGGCCC	1140
Dd		2634	CAGAAATGATCTTTGGAGAGATGTGTTCATTCTGTGAGTACTACAGCGTCTTTGACAGGCCC	2693
OY		1141	AACCACTCGTTGGGCTTAGCTAAAGCAATCTGA	1173
Dd		2694	AACCACTCGTTGGGCTTAGCTAAAGCAATCTGA	2726
RESULT 3				
AAN91157 standard; DNA; 1291 BP.				
XX	AC	AAN91157;		
XX	DT	10-MAR-2003 (updated)		
XX	DT	07-JUN-1990 (first entry)		
XX	DE	Cloned sequence of (pro)chymosin.		
XX	KW	Lactic acid bacteria; cheese; Streptococcus cremoris SK112; chymosin;		
XX	OS	proteinase; pSK112; ss.		
XX	FT	Unidentified.		
XX	Key	Location/Qualifiers		
XX	FT	precursor_RNA		
XX	FT	/tag= a		
XX	FN	/product=:prochymosin		
XX	FN	NL6701378-A.		


```

QY 56 TCGTTGCTGTGTAATGACGCGTGTGAGATGACCCGCAATTCCTCTCTACAAAGTAAGTCTC 115
DB 100 TCTTGCTCTCTCTCCAGGCGCGCTGAGATACAGAGATCCCTCTGTACAAAGGCAATCTCTC 159
QY 116 TCCGTAAGCGCGCTGAAGGAACATGGAATCTTAGAAGACTTCTTGCAAAAACAAGTAATG 175
DB 160 TGAGGAAGCGCGTGAAGGAACATGGCTCTTGAGAGACTTCCGCAAGAAACAGCACTAATG 219
QY 176 GCATCAGCAGCAAGTACTCCGCTTCCGTGAAGTTGCTAGCGTGCACCTTACCACTAC 235
DB 220 GCATCAGCAGCAAGTACTCCGCTTCCGTGAAGTTGCTAGCGTGCACCGTGCACCACTAC 279
QY 236 TTGATGATGAATACTTTGGGAAGATCTACTCGGAACCCCGCTCAAGAGTTACCGTTTC 295
DB 280 TGATAGTAGTACTTTGGGAAGATCTACTCGGAACCCCGCTCAAGAGTTACCGTTTC 339
QY 296 TCTTTGATAGTGGTCTCTGACTTCTGGGTTCCCTCTATCTACTGCAAGAGTAAGCTC 355
DB 340 TGTTTGACACTGGCTCTCTGACTTCTGGGTTCCCTCTATCTACTGCAAGAGTAAGCTC 399
QY 356 GCAGAAGCAACCAAGATTGATCCGAGAAAGTGTCCACCTTCCAGAACTTAGGCAAC 415
DB 400 GCAAAACCAACGCGCTTCCGAGAAAGTGTCCACCTTCCAGAACTTAGGCAACG 459
QY 416 CCTTGTCTATACCTACGCTACAGGTAGTACGCAAGAAATCTTAGGCTATGATACCGTCA 475
DB 460 CCTTGTCTATACCTACGCTACAGGTAGTACGCAAGAAATCTTAGGCTATGATACCGTCA 519
QY 476 CTGTCTCCAACATTTGACATTCACACAGACTAGGACTTAGACCAAGAACAGATG 535
DB 520 CTGTCTCCAACATTTGACATTCACAGAGACATGAGGCTTAGACCAAGAACAGATG 579
QY 536 ATGTCTTCAACCTATGAGAAATTCGATGGCATCTTGATGCAATCCATGCTGCGCT 595
DB 580 AGCTCTTCAACCTATGAGAAATTCGAGGATCTTGAGGAGATGCGCTACCTCGCTGCGCT 639
QY 596 CAGAGTACTGAGTACTGCTGTTTGAACAACATGATGAACCAACCTAGTACTGCAAGCT 655
DB 640 CAGAGTACTGAGTACTGCTGTTTGAACAACATGATGAACCAACCTAGTACTGCAAGCT 699
QY 656 TGTTCCTGCTTACATGACAGGAATGGCCAGAGAGCATGCTACGCTTGAAGTATTG 715
DB 700 TGTTCCTGCTTACATGACAGGAATGGCCAGAGAGCATGCTACGCTTGAAGTATTG 759
QY 716 ATTCATCTTACTACACAGATCTCTTCACTGGGTTCCAGTCTGTCAGCACTGCTG 775
DB 760 ACCGCTCTTACTACACAGAGGCTCCCTGCACTGGGTTCCGTCAGTCTGTCAGCACTGCTG 819
QY 776 AATTCACTGTGACAGTGTACATCAACGCGGTGTGTGTTGATGATGAAGGTGATGTC 835
DB 820 AGTTCACTGTGACAGTGTACATCAACGCGGTGTGTGTTGATGATGAAGGTGATGTC 879
QY 836 AAGCTTACTTGAATACCGGTAGCGTCAAGCTGGTCGGAAGCTTAGCAGGACATTTCTCAACA 895
DB 880 AAGCCATCTCTGACACGGGCACTTCAAGCTGGTCGGAAGCTTAGCAGGACATTTCTCAACA 939
QY 896 TTGAGCAAGCTATTTGAGCCACAGAAACAGTACGCTGATTTGACATGATTTGCGACA 955
DB 940 TTGAGCAAGCTATTTGAGCCACAGAAACAGTACGATGATTTGACATGATTTGCGAGACA 999
QY 956 ACTTTAGCTACATGCTTCACTGTTCTTTGATGATCAACGGCAAGATATCCCACTGACCC 1015
DB 1000 ACTTGAGCTACATGCTTCACTGTTCTTTGATGATCAATGCAAAATTTACCCCACTGACCC 1059
QY 1016 CCGTCCGCTATACAGCAGGCAAGGCTTTCGACACAGTGGATTCAGAGTGAAGAAC 1075
DB 1060 CCGTCCGCTATACAGCAGGCAAGGCTTTCGACACAGTGGCTTCAAGTGAAGAAATC 1119
QY 1076 ATTCCAGAAATGATCTTGGAGATGTTTCAATTCGTGATGATCAAGCGTCTTTGACA 1135
DB 1120 ATTCCAGAAATGATCTTGGAGATGTTTCAATTCGTGATGATTAAGAGTCTTTGACA 1179
QY 1136 GGGCCAAACAACCTGCTTGGGCTAGCTAAGCAATCTGA 1173

```

```

DB 1180 GGGCCAAACAACCTGCTTGGGCTAGCTAAGCAATCTGA 1217

RESULT 6
AAN40180
ID AAN40180 standard; cDNA; 1460 BP.
XX
AC AAN40180;
XX
AC 25-MAR-2003 (updated)
DT 25-JAN-1992 (first entry)
XX
DE Sequence of recombinant GFP4 carrying the remain coding sequence.
XX
KW Yeast expression vector; GAL1 promoter; Saccharomyces cerevisiae;
KW ss.
XX
OS Bos taurus.
XX
FH Key Location/Qualifiers
FT CDS 205..1350
FT /tag=a
XX
PN GB2137208-A.
XX
PD 03-OCT-1984.
XX
PF 28-FEB-1984; 84GB-0005129.
PR 28-FEB-1983; 8JUS-0470911.
XX
PA (COLB ) COLLABORATIVE RES INC.
XX
PI Botstein D, Davis RW, Fink GR, Tauntonig A, Knowlton RG, Mao JI,
PI Moir DT, Goff CG;
XX
DR WPI; 1984-245517/40.
DR P-PSDB; AAP40218.
XX
PT DNA segment contg. GAL1 promoter linked to gene - useful for
PT direction of expression of the gene in yeast cell
XX
PS Example; Table 4, Page 21-23; 35pp; English.
XX
CC The inventors claim a DNA segment contg. GAL1 promoter linked to
CC gene - useful for direction of expression of the gene in yeast cell.
CC The recombinant material carrying a GAL1 promoter of the yeast
CC galactokinase gene may be used in expressing a desired protein, esp.
CC bovine growth hormone, interferon, prorenin or preprorenin, in the
CC yeast cell. Strains of Saccharomyces cerevisiae producing the
CC polypeptides are produced. Yeast strains deposited as ATCC 20643,
CC 20661, 20662 and 20663, strain designations CGY 196, 457, 461 and
CC 528, resp. are new.
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 1460 BP; 328 A; 440 C; 399 G; 293 T; 0 other;

Query Match 76.5%; Score 897.2; DB 5; Length 1460;
Best Local Similarity 87.7%; Pred. No. 4e-273; Indels 0; Gaps 0;
Matches 980; Conservative 0; Mismatches 138;
QY 56 TCGTTGCTGTGTAATGACGCGTGTGAGATGACCCGCAATTCCTCTCTACAAAGTAAGTCTC 115
DB 233 TCTTGCTCTCTCTCCAGGCGCGCTGAGATACAGAGATCCCTCTGTACAAAGGCAATCTCTC 292
QY 116 TCCGTAAGCGCGCTGAAGGAACATGGAATCTTAGAAGACTTCTTGCAAAAACAAGTAATG 175
DB 293 TGAGGAAGCGCGTGAAGGAACATGGCTCTTGAGAGACTTCCGCAAGAAACAGCACTAATG 352
QY 176 GCATCAGCAGCAAGTACTCCGCTTCCGTGAAGTTGCTAGCGTGCACCGTGCACCACTAC 235

```

```

Db 353 GCATCAGACGACGATCTCCGGCTTCGGGGAGTGGCCAGCGTCCCTGACCAATCTAC 412
Oy 236 TTGATAGTCAATCTTTGGGAAGATCTACCTCGAAACCCCGCTGAGATTACCGTTTC 295
Db 413 TGGATAGTCAATCTTTGGGAAGATCTACCTCGAAACCCCGCTGAGATTACCGTTTC 472
Oy 296 TCTTTGATAGTCTGTTCTCTGACCTTCTGGGTTTCCCTCTATCTACTGCAAGCAATGCTT 355
Db 473 TGTTTGACATCTGGCTCTCTGACCTTCTGGGTTTCCCTCTATCTACTGCAAGCAATGCTT 532
Oy 356 GCAAGAACCAACCAAGATTCCATCCGAGAAAGTCTCCACTTTCAAGAACTTTAGCCAAAC 415
Db 533 GCAAAACACACAGGCTTCGACCCGAGAAATCTCTCCACTTTCAAGAACTTTAGCCAAAC 592
Oy 416 CCTTGTCTATACACTAGGTTACAGGTAGATGCAAGAACTCTTAGGCTATGATACCTGCA 475
Db 593 CCTGTCTATACACTAGGTTACAGGTAGATGCAAGAACTCTTAGGCTATGATACCTGCA 652
Oy 476 CTGTCTCAACACTTGTGACATTTCAACAGACAGTAGAATTAGCAACCAAGAACAGGTG 535
Db 653 CTGTCTCAACACTTGTGACATTTCAACAGACAGTAGAATTAGCAACCAAGAACAGGTG 712
Oy 536 ATGTCTTCACTATGACAGAAATTCGATGCGATCTTGTGATGSCATACCCATCTCGCT 595
Db 713 ACGTCTTCACTATGCGCAATTCGACGGGATCTGGGGATGCGCTACCCCTGCGCGCT 772
Oy 596 CAGAGTACTCGATACCTGTTTGAACAATGATAGACCACTAGTACTGCAAGT 655
Db 773 CAGAGTACTCGATACCTGTTTGAACAATGATAGACCACTAGTACTGCAAGT 832
Oy 656 TGTTCCTGGTTTACATGACAGAAATGCGCAGAGAGAGATGCTCAGCTTGGAGCTATTG 715
Db 833 TGTTCCTGGTTTACATGACAGAAATGCGCAGAGAGAGATGCTCAGCTTGGAGCTATTG 892
Oy 716 ATCATCTTACTACACAGAGATCTCTTCACTGSGTTTCACTGTCAGTGTGACAGTATG 775
Db 893 ACCGCTCTTACTACACAGAGATCTCTTCACTGSGTTTCACTGTCAGTGTGACAGTATG 952
Oy 776 AATTACATGTCAGTGTGTCATCAGATCAGCGGTGTTGTTGTCATGTCAGTGTGTCATG 835
Db 953 AGTTTACTGTGACAGTGTGTCATCAGATCAGCGGTGTTGTTGTCATGTCAGTGTGTCATG 1012
Oy 836 AAGTATCTTGGATACCGGTACGTGTCAGATGTCGAGACCTGACAGCGACATTTTCAACA 895
Db 1013 AGGCAATCTGTGACACGGGACCTTCAAGCTGTGCGGCGCCAGCAGACATCTTCAACA 1072
Oy 896 TTCAGCAAGCTATTGAGCCACACAGAACCACTACGTTGATGATGATGTCGACA 955
Db 1073 TCAGCAGAGCCATTTGAGCCACACAGAACCACTACGTTGATGATGATGTCGACA 1132
Oy 956 ACCTTAGCTACATGCTTACAGTGTGTTTGGATCAACGGCAAGTGTACCACTGACCC 1015
Db 1133 ACCTGAGGTACATGCTTACAGTGTGTTTGGATCAACGGCAAGTGTACCACTGACCC 1192
Oy 1016 CCTCGCCCTATACAGCCAGACGATCAAGGGTTCTGACACAGTGGATTCCAGAGTGAAGC 1075
Db 1193 CCTCGCCCTATACAGCCAGACGATCAAGGGTTCTGACACAGTGGATTCCAGAGTGAAGC 1252
Oy 1076 AATCCAGAAATGATCTTGGAGATGTTGTCATCTGAGTACTACAGCGTCTTTGACA 1135
Db 1253 AATCCAGAAATGATCTTGGAGATGTTGTCATCTGAGTACTACAGCGTCTTTGACA 1312
Oy 1136 GGGCCCAACACTCTGTTGGGCTAGCTAAAGCAATCTGA 1173
Db 1313 GGGCCCAACACTCTGTTGGGCTAGCTAAAGCAATCTGA 1350

```

RESULT 7
 AAN40055
 ID AAN40055 standard; DNA; 1278 BP.
 AC
 AAN40055;
 XX

```

DT 02-FEB-1992 (first entry)
XX
DE Sequence of prochimysin gene.
XX
KM Prochimysin expression vector; E.coli trp operon; chimysin; ss.
XX
OS Bos taurus.
XX
FH Key Location/Qualifiers
FT CDS 3..1130
FT polyA_signal /*tag= a
FT 1245..1250 /*tag= b
XX
PN EP121775-A.
XX
PD 17-OCT-1984.
XX
PF 07-MAR-1984; 84EP-0102451.
XX
PR 09-MAR-1983; 83JP-0038439.
XX
PA (BEPP/) BEPPU T.
XX
PI Beppu T, Uozumi T, Nishimori K, Shimizu N, Kawaguchi Y;
XX
PI Hidaka M;
XX
DR WPI; 1984-258001/42.
XX
DR P-PSDB; AAP40078.
XX
PT Expression plasmid comprising prochimysin gene and vector -
PT useful for transforming Escherichia coli for prochimysin prodn.
XX
PS Disclosure; Fig 1; 59pp; English.
XX
CC The inventors claim the prochimysin gene comprising a nucleotide
CC sequence from (a) the 1st codon (GCT) to the 365th codon (ATC), or
CC (b) the 5th codon (GGG) to the 365th codon (ATC); and recombinant
CC plasmids harbored by Escherichia coli strains deposited as FERM BP-
CC 262, -263 and -264. Any portion of the nucleotide sequence as
CC described in AAN40055 can be used. Also claimed is a vector derived
CC from plasmid pBR322. Typically plasmid pCR501 is obt'd. from pOCT 2.
CC The transcriptional direction of pOCT 3 is opposite to that of
CC pOCT 2; it is clockwise in pOCT 2 (5' to 3') whereas counter
CC clockwise in pOCT 3.
XX
SQ Sequence 1278 BP; 309 A; 384 C; 326 G; 259 T; 0 other;
XX
Query Match 76.4%; Score 895.6; DB 5; Length 1278;
Beet Local Similarity 87.6%; Pred. No. 1.2e-272;
Matches 979; Conservative 0; Mismatches 139; Indels 0; Gaps 0;
Oy 56 TCGTTGCTGTACTACACGCTGTCAGATCACCCGATCTCTCTACAAAGTATGCTC 115
Db 13 TCTTCGCTCTCTCCAGGCGCTGAGATCACAGATCCCTCTGTACAAAGCAATGCTC 72
Oy 116 TCCGTAGCGCCCTGAAGGACATGACCTTCTAGAACTCTTTCGAAACACAGATG 175
Db 73 TGAGGAAGCCCTGAAGGACATGAGCTTCTGAGAGACTTCTCGAAGAACAGCAGATG 132
Oy 176 GCATCAGACGACGATCTCGGCTTCCGTTGAGTGTGTAAGGTGCACTTACCAACTACC 235
Db 133 GCATCAGACGACGATCTCGGCTTCCGTTGAGTGTGTAAGGTGCACTTACCAACTACC 192
Oy 236 TTGATAGTCAATCTTTGGGAAGATCTACCTCGAAACCCCGCTGAGATTACCGTTTC 295
Db 193 TGGATAGTCAATCTTTGGGAAGATCTACCTCGGAAACCCCGCTGAGATTACCGTTTC 252
Oy 296 TCTTTGATAGTCTGTTCTCTGACCTTCTGGGTTTCCCTCTATCTACTGCAAGCAATGCTT 355
Db 253 TGTTTGACATCTGGCTCTCTGACCTTCTGGGTTTCCCTCTATCTACTGCAAGCAATGCTT 312
Oy 356 GCAAGAACCAACCAAGATTCCATCCGAGAAAGTCTCCACTTCCAGAACTTAGGCAAC 415

```

```

Db      313 GCAAAAACCAACGCGCTTTCACCCGAGAAAGTGTCCACCTTCAGAACTGGGAAAGC 372
Qy      416 CCTTGTCTATACACTACGCTACAGATGACATGCAAGAACTTAAAGCTATGATACCGTCA 475
Db      373 CCTGTCTATACACTACGCTACAGATGACATGCAAGAACTTAAAGCTATGATACCGTCA 432
Qy      476 CTGTCTCCCAACATTTGGAGCATTTCAACAGACATGAGACTTACCAACCAAGAACAGGTG 535
Db      433 CTGTCTCCCAACATTTGGAGCATTTCAACAGACATGAGACTTACCAACCAAGAACAGGTG 492
Qy      536 ATGTCTTCACTTACGAGATTTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 595
Db      493 AGCTTTTCACTTACGAGATTTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 552
Qy      596 CAGAGTACTGATACCTGTGTTTGAACAATGATGAAACCACTTATGATGATGATGATGATG 655
Db      553 CAGAGTACTGATACCTGTGTTTGAACAATGATGAAACCACTTATGATGATGATGATGATG 612
Qy      656 TGTTCCTGGTTTACATGAGACAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 715
Db      613 TGTTCCTGGTTTACATGAGACAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 672
Qy      716 ATCCATCTTACTACAGAGATCTCTTCACTGGGTTTCAAGTCACTGTGACAGATGATGAG 775
Db      673 ACCGCTCTTACTACAGAGATCTCTTCACTGGGTTTCAAGTCACTGTGACAGATGATGAG 732
Qy      776 AATTCATCTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 835
Db      733 AGTTCACTGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 792
Qy      836 AAGCTATCTTGGATACCGGATACCTGATGATGATGATGATGATGATGATGATGATGATG 895
Db      793 AGGCAATCTTGGATACCGGATACCTGATGATGATGATGATGATGATGATGATGATGATG 852
Qy      896 TTAGCAGAGCTTATGAGACCAACAGAACAGTACGAGTATGATGATGATGATGATGATG 955
Db      853 TTAGCAGAGCTTATGAGACCAACAGAACAGTACGAGTATGATGATGATGATGATGATG 912
Qy      956 ACCTTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1015
Db      913 ACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 972
Qy      1016 CTTTCCGCTTATGAGACCAACAGAACAGTACGAGTATGATGATGATGATGATGATG 1075
Db      973 CTTTCCGCTTATGAGACCAACAGAACAGTACGAGTATGATGATGATGATGATGATG 1032
Qy      1076 ATTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1135
Db      1033 ATTCCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1092
Qy      1136 GGGGCAACCACTGTTGGGCTTAAAGCAATCTGA 1173
Db      1093 GGGGCAACCACTGTTGGGCTTAAAGCAATCTGA 1130

```

RESULT 8

AAN20043
ID AAN20043 standard; DNA; 1460 BP.

XX AAN20043;

XX 16-DEC-1992 (first entry)

DE Pre-prorennin-A gene DNA sequence.

XX Pre-pro-rennin; rennin; prorennin; enzyme; EC-3.4.23.4; chymosin;

KW protease; milk-clotting enzyme; ss.

XX Bos taurus.

XX Key Location/Qualifiers

FT CDS 205..1350

```

FT      /tag= a
PT      /label=
DB      /note= "pre-prorennin-A gene"
GB2091271-A.
28-JUL-1982.
15-JAN-1982; 82GB-0001120.
01-DEC-1981; 81US-0325481.
16-JAN-1981; 81US-0225717.
(COLB ) COLLABORATIVE RES INC.
Alford BL, Mao J, Moir DT;
WPI; 1982-62028E/30 (62028E).
P-PSDB; AAP20038.
Transformed cells producing rennin and its precursors - contg.
PT appropriate recombinant DNA material
PS Disclosure; Table 1; 39pp; English.
CC Bases 1-204 and 1351 to 1460 are attached to the pre-prorennin but
CC can be removed and are not essential to use of the gene in
CC expression. The gene may be ligated into plasmid pCGE21 and
CC expressed in E. coli. The resulting expressed enzyme is a well
CC known milk-clotting enzyme used in cheese-making.
SQ Sequence 1460 BP; 329 A; 440 C; 397 G; 294 T; 0 other;
Query Match 76.4%; Score 895.6; DB 3; Length 1460;
Best Local Similarity 87.6%; Pred. No. 1.3e-272;
Matches 979; Conservative 0; Mismatches 139; Indels 0; Gaps 0;
Qy      56 TCGTGTCTGTTACTACGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 115
Db      233 TCTTCCCTCTCTCCGAGGCGCTGATGATGATGATGATGATGATGATGATGATGATG 292
Qy      116 TCGTAAAGGCGCGAAGAAACATGACTTCTGAAGACTTCTTGAAGAAACAGATATG 175
Db      293 TGAAGAAAGGCGCGAAGAAACATGACTTCTGAAGACTTCTTGAAGAAACAGATATG 352
Qy      176 GCATCAGCAGCAAGTATCTCGGCTTGTGAGATGATGATGATGATGATGATGATGATGATG 235
Db      353 GCATCAGCAGCAAGTATCTCGGCTTGTGAGATGATGATGATGATGATGATGATGATGATG 412
Qy      236 TTGATAGTCANATCTTTGGAGATCTTACCTGGAAACCCGCTCAAGAGTTCAACGTTT 295
Db      413 TGATAGTCAATCTTTGGAGATCTTACCTGGAGACCCGCTCAAGAGTTCAACGTTT 472
Qy      296 TCTTATATCTGTTCTCTGATCTTCTGATGATGATGATGATGATGATGATGATGATGATG 355
Db      473 TGTTCACATGCTCTCTGATCTTCTGATGATGATGATGATGATGATGATGATGATGATG 532
Qy      356 GCAAGAACCAACAAATTCGATCCGAGAAAGTGTGCACTTTCAGAACTTATGAGCAAC 415
Db      533 GCAAAAACCAACGCGCTTGCACCGAAGATGATGATGATGATGATGATGATGATGATGATG 592
Qy      416 CTTTGTCTATACACTACGCTACAGATGACATGCAAGAACTTAAAGCTATGATACCGTCA 475
Db      593 CCTGTCTATACACTACGCTACAGATGACATGCAAGAACTTAAAGCTATGATACCGTCA 652
Qy      476 CTGTCTCCCAACATTTGGAGCATTTCAACAGACATGAGACTTACCAACCAAGAACAGGTG 535
Db      653 CTGTCTCCCAACATTTGGAGCATTTCAACAGACATGAGACTTACCAACCAAGAACAGGTG 712
Qy      536 ATGTCTTCACTTACGAGATTTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 595
Db      713 AGCTTTTCACTTACGAGATTTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 772

```

QY 556 CAGAGTACTGATACCTGTGTTGACAAATGATGAACCGACCACTAGTACTCAAGACT 655
 DB 773 CAGAGTACTGATACCTGTGTTGACAAATGATGAACCGACCACTAGTACTCAAGACT 832
 QY 656 TGTTCCTGGTTTACATGACAGAAATGGCCAGAGACATGCTCAAGCTTGAAGCTATTG 715
 DB 833 TGTTCCTGGTTTACATGACAGAAATGGCCAGAGACATGCTCAAGCTTGAAGCTATTG 892
 QY 716 ATCCATCTACTACACAGAGATCTCTTCACTGGGTTCCAGTCACTGAGACAGATCTGGC 775
 DB 893 ACCCTGCTACTACACAGAGATCTCTTCACTGGGTTCCAGTCACTGAGACAGATCTGGC 952
 QY 776 AATTCACTGTGACAGATGTCACCAATCAAGCGGTGTGGTTGTTCCATGTCAGAGTGAATGTC 835
 DB 953 AGTTCACGTGTGACAGATGTCACCAATCAAGCGGTGTGGTTGTTCCATGTCAGAGTGAATGTC 1012
 QY 836 AAGCTATCTTGATACCGGTACGTCACAGCTGTCGAGCTTGAACGACGACATTTCTCAACA 895
 DB 1013 AGGCATCTGTGACACGCGACCTCCAAAGCTGTGCGGCGCCAGACGACATCTCAACA 1072
 QY 896 TTCAAGAGCTATTTGAGGACACAGAAACCACTAGCGTGAATTTGACATGATGTCGACCA 955
 DB 1073 TCCAGAGGCGCATTTGAGGACACAGAAACCACTAGCGTGAATTTGACATGATGTCGACCA 1132
 QY 956 ACCTTAGCTACATGCTTACAGTTGTTCTTGAATCAACGCGCAAGATGTACCCATGACCC 1015
 DB 1133 ACCTAGAGCTACATGCTTACAGTTGTTCTTGAATCAACGCGCAAGATGTACCCATGACCC 1192
 QY 1016 CCTCGGCTTATACCAAGCAGATCAAGGTTTGCACCACTGATTCAGAGTGAAGACC 1075
 DB 1193 CCTCGGCTTATACCAAGCAGATCAAGGTTTGCACCACTGATTCAGAGTGAAGACC 1252
 QY 1076 ATTCCAGAAATGATCTTGGAGATGTTGATTCATGTCGATGATCAACGCGCTTTGACA 1135
 DB 1253 ATTCCAGAAATGATCTTGGAGATGTTGATTCATGTCGATGATCAACGCGCTTTGACA 1312
 QY 1136 GGGCCAAACACTCGTGGGCTAGCTAAAGCAATCTGA 1173
 DB 1313 GGGCCAAACACTCGTGGGCTAGCTAAAGCAATCTGA 1350

RESULT 9
 ID AAZ06463
 AAZ06463 standard; DNA; 2733 BP.

XX AAZ06463;

DT 29-NOV-1999 (first entry)

DE 2.7 Kbp HindIII fragment of pSBSPORTNT (oleosin-chymosin fusion gene).

KW oil-body; lipid body; oleosome; spherosome; separation;
 fusion protein; heterologous polypeptide; commercial production;
 plasmid; ds.

OS Synthetic.

XX Key Location/Qualifiers

FT CDS 850..2729

FT /tag= a "oleosin-spacer-Met-prochymosin"

FT /note= "CDS contains an intron"

FT /number= 1

FT /number= 1

FT /number= 1

FT /number= 1

FT /number= 1

PD 07-SEP-1999.

XX US5948682-A.

XX 07-SEP-1999.

PF 25-APR-1997; 97US-0846021.
 PR 25-APR-1997; 97US-0846021.
 PR 22-FEB-1991; 91US-0659835.
 PR 16-NOV-1993; 93US-0142418.
 PR 30-DEC-1994; 94US-0366783.
 PA (SEMB-) SEMBIOSYS GENETICS INC.
 PI Moloney MM;
 DR WPI; 1999-517960/43.
 DR P-PSDB; AAY33830.
 PT Expression of a heterologous polypeptide on an oil body protein is
 useful for the production of e.g. enzymes, antibodies, hormones
 Claim 15; Fig 6; 48pp; English.
 CC This is the nucleotide sequence of a HindIII fragment containing the
 CC oleosin-spacer-Met-prochymosin sequence. This HindIII fragment was
 CC joined to a nopaline synthase terminator and cloned into binary vector
 CC pCEN1559. The resulting plasmid was called pSBSPORTNT and introduced
 CC into *A. tumefaciens*. The resulting bacterial strain was used to transform
 CC *B. napus* plants.
 CC The DNA which encodes a chimeric fusion protein that consists of the
 CC oil-body targeting sequence, a transcription regulation sequence and the
 CC DNA of the protein of interest can be used to produce antibodies,
 CC glycanases, hormones, proteases, protease inhibitors, seed storage
 CC proteins, thrombin inhibitors, hirudin, interleukins, chymosin, cystatin,
 CC xylanase, carp growth hormone, zein or a collagenase.
 CC The enzyme may be cleaved from the oil body protein or used in
 CC association with the oil body fraction.
 CC Allows production of commercially important proteins on a superior scale
 CC to production by conventional systems. The expressed heterologous
 CC protein can be easily separated from host cell components
 CC due to the use of the oil body as a carrier protein.
 SQ Sequence 2733 BP; 699 A; 681 C; 684 G; 669 T; 0 other;

Query Match 76.3%; Score 894.8; DB 20; Length 2733;
 Best Local Similarity 88.4%; Pred. No. 3.2e-272;
 Matches 971; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 76 GCTGAGATCACCCTGCTCTTCTACCAAGTATGTTCTTCTGTAAGCGCTGAAGACA 135
 DB 1630 GCTGAGATCACCCTGCTCTTCTGACAAAGGCAAGTCTTGAGGAAGCGCTGAAGAG 1689
 QY 136 CATGACTTCTAGAACTTCTTGAGAAACAGATGATGACAGCAAGTACTCC 195
 DB 1690 CATGAGCTTCTGAGAGATCTTCTGAGAAACAGATGATGACAGCAAGTACTCC 1749
 QY 196 GGTTCGTGGAAGTTGCTAGCGTGCACCTTACCAACTCTTGAATGATCAATTAATCTTGG 255
 DB 1750 GGTTCGTGGAAGTTGCTAGCGTGCACCTTACCAACTCTTGAATGATCAATTAATCTTGG 1809
 QY 256 AAGATCTACTCGAACCCTGCTCAGAGTTCACTGTTCTTTGATGATGTTCTCT 315
 DB 1810 AAGATCTACTCGAACCCTGCTCAGAGTTCACTGTTCTTTGATGATGTTCTCTCT 1869
 QY 316 GACTTCTGAGTCTCTCTATCTACGCAAGAGCATGCTCAGAAACACCAAGATTC 375
 DB 1870 GACTTCTGAGTCTCTCTATCTACGCAAGAGCATGCTCAGAAACACCAAGATTC 1929
 QY 376 GATCGAAGAAAGTGTCTCAGACCTTCCAGAACTTAGGCAAACTTGTCTATACATTCAGT 435
 DB 1930 GATCGAAGAAAGTGTCTCAGACCTTCCAGAACTTAGGCAAACTTGTCTATACATTCAGT 1989
 QY 436 ACAGTACGATGACAGAAATCTTAGGCTATGATACCTGCTATCTTCAATTTGAGAC 495
 DB 1990 ACAGTACGATGACAGAAATCTTAGGCTATGATACCTGCTATCTTCAATTTGAGAC 2049
 QY 496 ATTCAACAGACATGACATTAAGCAACCAAGAACAGAGTGTATGATCACTTATGACAGA 555

```

Db      2050 ATCCAGACAGACATAGGCTTGGAGACCCAGAGACCCGGGAGCTTTCACCTTATGCCGAA 2109
QY      556 TTGATGGACATCCTTGGTATGGCATACCCATCGCTCGCGTCAAGATAGTGATACCTGTG 615
Db      2110 TTGCAGCGGATCTCGGGATAGGCTTACCTCGCTGCGCTCAAGATCTGATACCCGTG 2169
QY      616 TTGACACATGATGAAACCGACACTAGTAGCTCAAGACTTGTCTCGGTTTACATGAGAC 675
Db      2170 TTGACACATGATGAAACCGACACTAGTAGCTCAAGACTTGTCTCGGTTTACATGAGAC 2229
QY      676 AGAATGGCCAGAGAGACATGCTACAGCTTGAAGCTTATGATCATCTCTACTACACAGA 735
Db      2230 AGAATGGCCAGAGAGACATGCTCTACGCTGGGGCCATTCACCCGTCTTACTACACAGG 2289
QY      736 TCTCTTCTACTGGTTCAGTCACTGTGACAGCAAGTACTGGCAATTCATGTGACAGTGTG 795
Db      2290 TCCCTGACTGGGTGCGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2349
QY      796 ACCATCAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 855
Db      2350 ACCATCAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2409
QY      856 ACGTCAAGCTGTGCGGACCTAGACGAGACATTCCTCAACATTCAGCAAGCTATTGAGACC 915
Db      2410 ACGTCAAGCTGTGCGGACCTAGACGAGACATTCCTCAACATTCAGCAAGCTATTGAGACC 2469
QY      916 ACAAGACAGTACCGGTGAGTTGACATAGATTGCGAACACTTATGCTACATGCTTACA 975
Db      2470 ACACAGAACAGTACCGGTGAGTTGACATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2529
QY      976 GTTGTCTTGAATGACAGGCAAGATATACCACTGACCCCTCGCTTATCCAGGACAG 1035
Db      2530 GTTGTCTTGAATGACAGGCAAGATATACCACTGACCCCTCGCTTATCCAGGACAG 2589
QY      1036 GATCAAGGGTTCGACACAGTGAATTCAGAGTGAAGAACATTCGCAAGATGGATCTTG 1095
Db      2590 GACCAAGGGTTCGACACAGTGAATTCAGAGTGAAGAACATTCGCAAGATGGATCTTG 2649
QY      1096 GGAATGTGTTCATTCGTGATGATCTACAGCGCTTTTGAAGGCGCAACACTGTGTGG 1155
Db      2650 GGGGATGTTCATTCGAGATATTAAGCGTCTTTTGAAGGCGCAACACTGTGTGG 2709
QY      1156 CTAGCTTAAGCATCTGA 1173
Db      2710 CTGGCCAAAGCCATCTGA 2727

RESULT 10
ABS53073
ID      ABS53073 standard; DNA; 2733 BP.
XX
AC      ABS53073;
XX
DT      29-NOV-2002 (first entry)
XX
DE      DNA encoding oleosin-chymosin fusion protein.
XX
KM      Thioredoxin; thioredoxin reductase; gene expression; oleosin;
XX      oil body; oleosin-chymosin fusion protein; gene; ds.
XX
OS      Arabidopsis thaliana.
XX
OS      Bos sp.
XX
OS      Synthetic.
XX
FH      Key
XX      CDS
XX      Location/Qualifiers
FT      /*tag= a
FT      /product= "Oleosin-chymosin fusion protein"
FT      exon
FT      /tag= b
FT      /number= 1
FT      intron
FT      1204..1440

```

```

FT      /*tag= c
FT      /number= 1
FT      exon
FT      1441..2727
FT      /*tag= d
FT      /number= 2
FT      misc_difference 1608..1630
FT      /*tag= e
FT      /note= "spacer sequence replacing the oleosin stop
FT      codon"
XX
XX      US2002088025-A1.
XX
XX      04-JUL-2002.
XX
XX      03-JUL-2001; 2001US-0897425.
XX
XX      22-FEB-1991; 91US-0659835.
XX      16-NOV-1993; 93US-0142418.
XX      30-DEC-1994; 94US-0366783.
XX      25-APR-1997; 97US-0846021.
XX      18-DEC-1998; 98US-0210843.
XX
XX      (MOLO/) MOLONEY M. M.
XX      (DALM/) DALMIA B. K.
XX
XX      Moloney MM, Dalmia BK;
XX
XX      WPI; 2002-635723/68.
XX      P-PFSD; ABG32904.
XX
XX      Expressing protein, by introducing chimeric nucleotide regulatory
XX      sequence, sequence encoding fusion protein, having sequence encoding
XX      protein, oleosin gene and sequence encoding termination region and
XX      producing protein -
XX
XX      Example 10; Fig 6; 69pp; English.
XX
XX      The invention describes a method of expressing thioredoxin or thioredoxin
XX      reductase (1) in the oil body of a host cell using an oil body protein
XX      gene. The method involves introducing a chimeric nucleic acid comprising
XX      a first sequence to regulate transcription, a second DNA sequence
XX      encoding a fusion polypeptide, comprising a sequence encoding an oleosin
XX      gene, and sequence encoding (1) and a third sequence encoding a
XX      termination region functional in the host cell and growing the host cell
XX      to produce a fusion polypeptide. The method or (1) is useful for
XX      expression of a thioredoxin or thioredoxin reductase by a host cell. This
XX      sequence encodes an Arabidopsis thaliana oleosin-bovine aspartic
XX      protease, chymosin fusion protein.
XX
XX      Sequence 2733 BP; 699 A; 681 C; 684 G; 669 T; 0 other;
XX
XX      Query Match
XX      Best Local Similarity 88.4%; Score 894.8; DB 24; Length 2733;
XX      Matches 971; Conservative 0; Mismatches 127; Indels 0; Gaps 0;
XX
QY      76 GCTGAGTACCCGCAATTCCTCTCTACAAAGTAACTCTCTCCGTAAAGCGCTGAAGAA 135
Db      1630 GCTGAGATCACCGATCCCTCTGTACAAAGCAATCTCTGTGGAAGGCGCTGAAGGAG 1689
QY      136 CANGAAGCTTGAAGACTTCTTGAGAAACAAGATAGGATCGACAGCAAGTACTCC 195
Db      1690 CATGGCTTCTGAGAGACTTCTGCAAGAAACAGCAATATGATCGACAGCAAGTACTCC 1749
QY      196 GAGCTGAGTAAAGTGTCTAGCGTGCACCTTACCACTTACCTTGAATGATCAATCTTTGG 255
Db      1750 GAGTTGGGAGAGTGGCAGCGCGTCCCTGACCAACTTACTGATAGTACTTTGG 1809
QY      256 AAGATCTACTCGGAACCCCGCTCAAGAGTTCACCGTCTCTTTGATAGTATGTTCTCT 315
Db      1810 AAGATCTACTCGGAACCCCGCTCAAGAGTTCACCGTCTCTTTGATAGTATGTTCTCT 1869
QY      316 GACTTCTGGGTTCCCTCTATCTACTGCAAGAGAAAGCTTGCAGAAAGCAACCAAGATT 375

```

```

Db 1870 GACTTCTGGGTACCTCTATCTACTGTGCAAGCAAAATGCTGCAAAAACCAACGAGCGCTTC 1929
Qy 376 GATCCGAGAAAGTCTGCTCACTTCCAGAACTTAGGCAAAACCTTGCTATACACTAGCT 435
Db 1930 GACCCGAGAAATCGTCCACTTCCAGAACTGGGCAAGCCCTGTCTATCCACTACAGG 1989
Qy 436 ACAGGTAGCATCAAGAAATCTTAGGCTATGATACCGTCTCTCCAACTTGTGAC 495
Db 1990 ACAGGAGCATCAAGGAGCTCTGGGCTATGACACCGTCACTGTCTCCAACTTGTGAC 2049
Qy 496 ATTCAACAGAGTGAAGTATGACACCCAGAAACCAAGTGAAGTCTTCCACTATGACAGA 555
Db 2050 ATCCAGCAAGAGTGAAGCTCTGAGACCCAGAGCCCGGGAAGCTTCACTATGCGGA 2109
Qy 556 TTCGATGGCATCCTTGATGAGCATACCATGCTCTGCGCTCAAGAGTACGATACCTGTG 615
Db 2110 TTCGACGGGATCCTGGGATGAGCTACCCCTGCTGCTCAAGAGTACGATACCCGTG 2169
Qy 616 TTTGACAACTATGAAACCGACACCTAGTACTCAAGACTTGTCTCGGTTTACATGAC 675
Db 2170 TTTGACAACTATGAAACCGACACCTAGTACTCAAGACTTGTCTCGGTTTACATGAC 2229
Qy 676 AGGAATGGCCAGAGAGCATGCTCAAGCTTGAAGCTATGATCATCTTACTACAGAGA 735
Db 2230 AGGAATGGCCAGAGAGCATGCTCAAGCTTGAAGCTATGATCATCTTACTACAGAGG 2289
Qy 736 TCTCTTCACTGGGTTCCAGTCACTGAGCAGTACTGAGCAATTCATCTGTGACAGTGC 795
Db 2290 TCTCTGCACTGGGTTCCAGTCACTGAGCAGTACTGAGCAATTCATCTGTGACAGTGC 2349
Qy 796 ACCATCAGCGGTGTGTTGTTCATGATGAAAGTGTGATGCAAGCTATCTTGTGATACCGGT 855
Db 2350 ACCATCAGCGGTGTGTTGTTCATGATGAAAGTGTGATGCAAGCTATCTTGTGATACCGGT 2409
Qy 856 AGCTCCAGCTGCTGAGCACTTGAAGCAAGCATTTCTCAACTTCAAGCACTTATGAGCC 915
Db 2410 AGCTCCAGCTGCTGAGCACTTGAAGCAAGCATTTCTCAACTTCAAGCACTTATGAGCC 2469
Qy 916 AACAGAAACCACTAGCGTATGATTTGACATAGATTTGGAGAACTTAGCTACATGCTTAA 975
Db 2470 AACAGAAACCACTAGCGTATGATTTGACATAGATTTGGAGAACTTAGCTACATGCTTAA 2529
Qy 976 GTTGTCTTTGAGATCAACGCGCAAGATGTACCACTGACCCCTCGCTATACAGAGCCAG 1035
Db 2530 GTTGTCTTTGAGATCAACGCGCAAGATGTACCACTGACCCCTCGCTATACAGAGCCAG 2589
Qy 1036 GATCAAGGGTCTGCAACAGTGGATTCAAGTGAAGAACCATTTCCAGAAATGATCTTG 1095
Db 2590 GATCAAGGGTCTGCAACAGTGGATTCAAGTGAAGAACCATTTCCAGAAATGATCTTG 2649
Qy 1096 GGAAGATGTCTTCACTTGTGATGATCTACAGCGTCTTTGACAGGGCCAAACCTCTTG 1155
Db 2650 GGAAGATGTCTTCACTTGTGATGATCTACAGCGTCTTTGACAGGGCCAAACCTCTTG 2709
Qy 1156 CTAGCTAAAGCACTTGA 1173
Db 2710 CTAGCTAAAGCACTTGA 2727

```

RESULT 11
AAQ14051
ID AAQ14051 standard; DNA; 1210 BP.

XX AAQ14051;
AC 25-MAR-2003 (updated)
DT 06-JAN-1992 (first entry)
XX
DE Remnin gene.
XX
KW Prorennin; alpha-Si-casein gene; insulin-like growth factor I; IGF-I;
XX mammary gland; ss.

```

OS Synthetic.
XX
XX EP451823-A.
XX
XX 16-OCT-1991.
XX
XX 10-APR-1991; 91EP-0105702.
XX
XX 11-APR-1990; 90DE-4012526.
XX
XX 11-APR-1990; 90DE-4011751.
XX
XX (CONE ) CONSORTIUM ELEKTROCHEM IND.
XX
XX Har1 P, Brem G;
XX
XX WPI; 1991-304858/42.
XX
XX Recombinant DNA constructs for expressing protein in milk -
XX PT contg. specific mammary gland transcription control region and
XX PT signal sequence, providing high yield and easy prod. recovery
XX
XX Discloure; Page 21; 41dp; German.
XX
XX The rennin gene was isolated using the probe represented in AAQ14775.
XX CC It was used as heterologous peptide/protein together with parts of the
XX CC alpha-Si-casein gene in the prodn. of DNA constructs. The heterologous
XX CC peptide or protein may also be human insulin-like growth factor I.
XX CC The constructs provide high yields of the protein with simple recovery
XX CC from the milk. Activation of the gene occurs only in the mammary gland.
XX CC See also AAQ14050, AAQ14774-77.
XX CC (updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1210 BP; 278 A; 356 G; 324 C; 252 T; 0 other;

Query Match 76.2%; Score 894; DB 12; Length 1210;
Best Local Similarity 87.5%; Pred. No. 3.7e-272;
Matches 978; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 56 TCGTGTCTGTACTACGCTGTGATCACCCGATTCCTCTACAAAGGTAGTCTC 115
Db 43 TTGTGGCTGTGTCTTTCGCTGATCAATCAAGATCCCTGTACAAAGGCAAGTCTC 102
Qy 116 TCCGTAAAGCGGTGAAGAAACATGACTTCTAGAACTTTCTGCAAGAAACAGATAG 175
Db 103 TGAGGAAGCGGTGAAGAGATGAGGCTTTGAGAGACTTCTGCAAGAAACAGAGATAG 162
Qy 176 GCATCAGCAGCAAGTACTCCGGCTTGGTGAAGTGTCTAGCGTGCATTAACCACTAC 235
Db 163 GCATCAGCAGCAAGTACTCCGGCTTGGTGAAGTGTCTAGCGTGCATTAACCACTAC 222
Qy 226 TTGATATGATCACTTTGGGAAAGATCTACCTCGGAAACCCCGCTCAAGATTCACGTTT 295
Db 223 TGGATATGATCACTTTGGGAAAGATCTACCTCGGAAACCCCGCTCAAGATTCACGTTT 282
Qy 296 TCTTGTATATGAGTCTCTGATCTTCTGAGGTTCCCTCTATCTACTGCAAGAGCAATGCT 355
Db 283 TCTTGTATATGAGTCTCTGATCTTCTGAGGTTCCCTCTATCTACTGCAAGAGCAATGCT 342
Qy 356 GCAAGAACCAACCAAGATTCGATCCGAGAAAGTGTCTCACTTCCAGAACTTAGGCAAC 415
Db 343 GCAAGAACCAACCAAGATTCGATCCGAGAAAGTGTCTCACTTCCAGAACTTAGGCAAC 402
Qy 416 CTTGTCTATATCACTAAGGTATGATGATGATGATGATGATGATGATGATGATGATGAT 475
Db 403 CTTGTCTATATCACTAAGGTATGATGATGATGATGATGATGATGATGATGATGATGAT 462
Qy 476 CTGTCTCAACATTTGTGACATTTCAACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 535
Db 463 CTGTCTCAACATTTGTGACATTTCAACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 522
Qy 536 ATGTCTTCACTTATGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 595
Db 523 ACGTCTTCACTTATGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 582

```

QY 596 CAGAGTACTGATGACTGTGTGTTGACAACTGATGAACCGACCTTAGTACTGACACT 655
XX
Db 583 CAGAGTACTGATGACTGTGTGTTGACAACTGATGAACCGACCTGTGTGCCAAGACC 642
QY 656 TGTTCGAGTTTACATGAGCAGGAATGGCCAGAGAGACATGCTCAGGCTTGGAGCTATTG 715
XX
Db 643 TGTTCGAGTTTACATGAGCAGGAATGGCCAGAGAGACATGCTCAGGCTTGGAGCTATTG 702
QY 716 ATCCATCTCTACTACAGAGATCTCTTCACTGGGTTCCAGTCACTGTGACAGAGTACTGGC 775
Db 703 ACCGCTCTACTACAGAGATCTCTTCACTGGGTTCCAGTCACTGTGACAGAGTACTGGC 762
QY 776 AATTCACTGTGACAGTGTACCATCAGCGGTGTGTGTTGATGATGAAAGTGTATGTC 835
Db 763 AGTTCACTGTGACAGTGTACCATCAGCGGTGTGTGTTGATGATGAAAGTGTATGTC 822
QY 836 AACCTTCTTGGATACCGGTACGTCAGGCTGGTCCAGCTAGCAGAGACATTCCTCAACA 895
Db 823 AGGCACTCTGAGACAGGACCTCCAGGCTGGTCCAGGCTAGCAGATATCTCAACA 882
QY 896 TTCAGCAAGCTATTGAGCCACACAGAACAGTACGGTGTGATGATGATGATGATGATG 955
Db 883 TTCAGCAAGCTATTGAGCCACACAGAACAGTACGGTGTGATGATGATGATGATGATG 942
QY 956 ACCTTAGCTACATGCTCTACATGCTTCTTGAATGAAAGGAGATGATGATGATGATGATG 1015
Db 943 ACCTTAGCTACATGCTCTACATGCTTCTTGAATGAAAGTATGATGATGATGATGATG 1002
QY 1016 CCTCCGCTTACAGCAGGACGATCAAGGTTCTGACAGTGTGATGATGATGATGATGATG 1075
Db 1003 CCTCCGCTTACAGCAGGACGATCAAGGTTCTGACAGTGTGATGATGATGATGATGATG 1062
QY 1076 ATTCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1135
Db 1063 ATTCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1122
QY 1136 GGGCCAAACACCTGCTGGGTAGCTAAAGCAATCTGA 1173
Db 1123 GGGCCAAACACCTGCTGGGTAGCTAAAGCAATCTGA 1160
RESULT 12
AAN30209
ID AAN30209 standard; cDNA; 1290 BP.
XX
AC AAN30209;
XX
DT 25-MAR-2003 (updated)
DT 03-AUG-1992 (first entry)
XX
DE Sequence of preprochymosin cDNA.
XX
KM Milk-clotting; cheese making; enzyme; zymogen; ss.
XX
OS Cow.
XX
FH Key Location/Qualifiers
FT CDS 26..73
FT CDS /*tag= a
FT CDS /*tag= b
FT CDS /*tag= c
FT CDS /*tag= chymosin
XX
GB2100737-A.
XX
PD 06-JAN-1983.
XX
PF 11-JUN-1982; 82GB-0017096.

XX
PR 11-JUN-1982; 82GB-0017096.
XX
PA (CLUT) CELLTECH LTD.
XX
PI Carey NH, Harris TUR, Lowe PA, Doel MT, Emtage JS;
XX
DR WPI; 1983-00545K/01.
XX
DR P-PSDB; AAP30446.
XX
PT Prod. or calf stomach chymosin for cheese making - by
XX
PS cultivation of micro-organisms transformed with vector system
XX
PS Claim 42; Fig 4; 26pp; English.
XX
PS The inventors claim a method for the prodn. of calf stomach chymosin
XX
CC for cheese making. Genes and polypeptides for preprochymosin,
XX
CC prochymosin and chymosin are claimed, as are vector systems and a
XX
CC prochymosin primer.
XX
CC (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 1290 BP; 309 A; 384 C; 334 G; 263 T; 0 other;
Query Match 76.2%; Score 894; DB 4; Length 1290;
Best Local Similarity 87.5%; Pred.No.3.8e-272;
Matches 978; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 56 TCCTGCTCTGTTACTACGCTGTGATGATCAACCGCATTCCTCTCAAAAGTATGCTC 115
Db 54 TCCTGCTCTCTCTCCCAAGGCGCTGATATCAACGATTCCTCTGTCAAAAGGATGCTC 113
QY 116 TCCGTAAGGCGCTGGAAGAAATGATGATCTTGAAGACTTCTTGAAGAAATGATG 175
Db 114 TGAAGAGGCGCTGGAAGAGATGAGGCTTCTGAGAGACTTCTTGAAGAAATGATG 173
QY 176 GCATCAGCAGCAAGTACTCCGCTTGGTGAAGTGTGCTGAGCGTCAAGTATCAACTAC 235
Db 174 GCATCAGCAGCAAGTACTCCGCTTGGGAGGTGCGAGCGTCAAGTATCAACTAC 233
QY 236 TTGATAGTCAATCTTTGGGAAGATCTACCTCGAAGCCCGCTCAAGATTCACCGTTC 295
Db 234 TAGATAGTCAATCTTTGGGAAGATCTACCTCGAAGCCCGCTCAAGATTCACCGTTC 293
QY 296 TCTTGTATATGCTGCTCTGATCTTGTGGGTTCCCTATATGATGATGATGATGATG 355
Db 294 TGTTCGACTGCTCTCTGATCTTGTGGGTTCCCTATATGATGATGATGATGATG 353
QY 356 GCAAGAACCAACCAAGATTCGATCCGAGAAAGTCTCACTTCCAGAACTTAGGCAAC 415
Db 354 GCAAGAACCAACCAAGATTCGATCCGAGAAAGTCTCACTTCCAGAACTTAGGCAAGC 413
QY 416 CCTGTCTATACCTACGAGTACGATGATGATGATGATGATGATGATGATGATGATG 475
Db 414 CCTGTCTATACCTACGAGTACGATGATGATGATGATGATGATGATGATGATGATG 473
QY 476 CTGTCTCAACATTTGAGCAATTCACACAGACATGATGATGATGATGATGATGATG 535
Db 474 CTGTCTCAACATTTGAGCAATTCACACAGACATGATGATGATGATGATGATGATG 533
QY 536 ATGTCTTCACTTATGAGAAATTTGATGATGATGATGATGATGATGATGATGATG 595
Db 534 ACCTCTTCACTTATGAGAAATTTGATGATGATGATGATGATGATGATGATGATG 593
QY 596 CAGAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 655
Db 594 CAGAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 653
QY 656 TGTTCGAGTTTACATGAGCAGGAATGGCCAGAGAGACATGCTCAGGCTTGGAGCTATTG 715
Db 654 TGTTCGAGTTTACATGAGCAGGAATGGCCAGAGAGACATGCTCAGGCTTGGAGCTATTG 713
QY 716 ATCCATCTCTACTACAGAGATCTCTTCACTGGGTTCCAGTCACTGTGACAGAGTACTGGC 775

Db 714 ACCCTCTACTACACAGGTCCTCCGCACTGGGTCCCGTGACAGTGCAGCTACTGCGC 773
Qy 776 AATTACTGTGGACAGTGTGCACATCAGCGGTGTGTGGTGCATGTGAAGGTGATGTC 835
Db 774 AGTTACTGTGACAGTGTGCACATCAGCGGTGTGTGGTGCATGTGAAGGTGATGTC 833
Qy 836 AAGCTATCTTGGATACCGGTACGTTCACAGCTGTGTGCACTTACAGCAGCAATTCACAA 895
Db 834 AGGCATCTTGGACAGCGGCACCTCCAGCTGTGTGGGCCCCAGCAGCAGCACTCCACAA 893
Qy 896 TTCAGCAAGCTATTGGAGCCACACAGAAACAGTACGGTGAATTGACATAGATTGCCA 955
Db 894 TCAGACAGCCATGTGAGCCACACAGAAACAGTACGGTGAATTGACATAGATTGCCA 953
Qy 956 ACCCTAGCTACATGCTCAGTGTCTTGGATCAAGCGCAAGATGTAACCCAGTACCC 1015
Db 954 ACCTAGCTACATGCTCAGTGTCTTGGATCAAGCGCAAGATGTAACCCAGTACCC 1013
Qy 1016 CCTCGCTATACAGCCAGATCAAGGTTCTGCACAGTGCATTCAGAGTGAAC 1075
Db 1014 CCTCGCTATACAGCCAGATCAAGGTTCTGCACAGTGCATTCAGAGTGAAC 1073
Qy 1076 ATTCCAGAAATGATCTTGGAGATGTTCATTGTGAGTATCAAGCGTCTTTGACA 1135
Db 1074 ATTCCAGAAATGATCTTGGAGATGTTCATTGTGAGTATCAAGCGTCTTTGACA 1133
Qy 1136 GGGCCAAACACTCGTGGGCTAGCTAAAGCAATCTGA 1173
Db 1134 GGGCCAAACACTCGTGGGCTAGCTAAAGCAATCTGA 1171

RESULT 13

AAAN91188 standard; DNA; 2727 BP.

AAAN91188;
25-MAR-2003 (updated)
15-JUN-1990 (first entry)
BamHI/Sali insert of Kluveromyces plasmid PAB309.
Kluveromyces; pDM100PC; chymosin; tissue plasminogen activator.
Kluveromyces lactis.
Key Location/Qualifiers
CDS 409..1781
/*tag= a
EP301670-A.
01-FEB-1989.
28-JUL-1988; 88BP-0201632.
28-JUL-1987; 87US-0078539.
(KONN) GIST-BROCADES NV.
Vandenbergh JA, Vanooijen AJU, Rietveld K;
WPI; 1989-033565/05.
P-PsDB; AAP94376.
Kluveromyces host cells for producing polypeptides) -
used for highly efficient prodn. of eg chymosin tissue
plasminogen activator or human serum albumin.
Disclosure; Page ?; 56pp; English.
BamHI/Sali insert, incorporated into plasmid PAB309 with

CC promoters, terminators and a G418 resistance marker fused to a ADHI
CC promoter from S.cerevisiae.
CC The Kluveromyces expression systems provide highly efficient secretion
CC and processing of a wide variety of proteins.
CC Sequences identical to those published in EP301669.
CC (Updated on 25-MAR-2003 to correct PD field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
SQ Sequence 2727 BP; 752 A; 654 C; 572 G; 749 T; 0 other;

Query Match 76.1%; Score 893; DB 10; Length 2727;
Best Local Similarity 87.5%; Pred. No. 1,2e-271;
Matches 977; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 57 CGTTCGTGTACTACCGCTGTGATGATACCCGATTCCTCTCAAAAGGTAAGTCTCT 116
Db 669 CGATGCTTCCATATATGCTTGAGATACAGATCCCTCTGACAAAGGCAAGTCTCT 728
Qy 117 CCGTAGCGCTGAAGAAACATGACCTTCTAGAACTTCTTGCAGAAACAGATATG 176
Db 729 GAGGAAGCGCTGAAGAGATGGGCTTCTGAGGACTTCTGCAAGAAACAGATATG 788
Qy 177 CATGACGACAGTATCTCCGCTTCTGGAAGTGTGACGTGCTTCACTTCACTACT 236
Db 789 CATGACGACAGTATCTCCGCTTCTGGAAGTGTGACGTGCTTCACTTCACTACT 236
Qy 227 TGATATGTAATCTTGGAGATCTACCTGGAACCCCGCTCAAGAGTTCAACGTTCT 296
Db 849 GAGACGTAGTACTTGGAGATCTTCTGGAACCCCGCTCAAGAGTTCAACGTTCT 908
Qy 297 CTTGATATGTTCTCTGACTTCTGAGTTCTGAGTTCTCTCTATCTACTGCAAGCAATGCTG 356
Db 909 GTTGACACTGCTCTCTGACTTCTGAGTTCTCTCTATCTACTGCAAGCAATGCTG 968
Qy 1029 CCTGCTATCCATCAAGGACAGGACAGATGAGGACCTCTGCTATGACACCGTCA 1088
Db 477 TGTCTCCAACTTGGACATTCACAGACAGTATGACCTTACGACCCCAAGAACAGTGA 536
Qy 1089 TGTCTCCAACTTGGACATTCACAGACAGTATGACCTTACGACCCCAAGAACAGTGA 1148
Qy 537 TGTCTCCAACTTGGACATTCACAGACAGTATGACCTTACGACCCCAAGAACAGTGA 596
Db 1149 GGTCTTCACTATGCGAATTCGACGGGATCTGCGGATGAGCTTACCTCGCTGCTC 1208
Qy 597 AGAGTACTGATACCTGTGTTGACACATGATGAAACCGACACTTATGACTCAAGACTT 656
Db 1209 AGAGTACTGATACCTGTGTTGACACATGATGAAACCGACACTTATGACTCAAGACTT 1268
Qy 657 GTTCTCGGTTTAACTGACAGAAATGCGCAGAGAGCATGCTGAGGATATTA 716
Db 1269 GTTCTCGGTTTAACTGACAGAAATGCGCAGAGAGCATGCTGAGGATATTA 1328
Qy 717 TCATCTCTACTACAGAGATCTTCTGAGTTCTCACTGACCTGACAGATATGCA 776
Db 1329 CCGCTCTACTACAGAGATCTTCTGAGTTCTCACTGACCTGACAGATATGCA 1388
Qy 777 ATTCACTGTGACAGTGTACCATGACGGGTGTGTTGTCATGTGAAGTGTGATGCA 836
Db 1389 GTTCACTGTGACAGTGTACCATGACGGGTGTGTTGTCATGTGAAGTGTGATGCA 1448
Qy 837 AGCATCTTGGATACCGGTATGCTCAAGCTGTGGAACCTTACAGAGCAATTTCAAT 896
Db 1449 GGCATCTTGGACAGGACCTTCAAGCTGTGGAACCTTACAGAGCAATTTCAAT 1508
Qy 897 TCAGCAAGCTATTGGAGCCACAGAAACAGTACGGTGAATTGACATAGATTGCCA 956
Db 1509 COAGAGCCATTTGGAGCCACAGAAACAGTACGGTGAATTGACATAGATTGCCA 1568

```
QY 957 CCTAGCTACATGCTCAGATGTTGCTTTGATCAACGCAAGATGTACCAGTCC 1016
DB 1569 CTTGAGCTACATGCCA CTGTGGCTTTTGATCAATGGCAAAATGTACCCAGTACC 1628
QY 1017 CTCGCGCTATACAGCAGAGATCAAGGGTCTGCACCAAGTGATTTCCAGATGAGAACCA 1076
DB 1629 CTCGCGCTATACAGCAGAGATCAAGGGCTTCTGTACAGAGTGGCTTCCAGATGAAATCA 1688
QY 1077 TTCCCAAGAAATGATCTTTGGAGATGTGTTTCATTCTGTAGTACTACAGCGCTTTGACAG 1136
DB 1689 TTCCCAAGAAATGATCTTTGGAGATGTGTTTCATTCTGTAGTACTACAGCGCTTTGACAG 1748
QY 1137 GGCCCAACACCTGTGGGCTAGCTAAAGCAATCTGA 1173
DB 1749 GGCCCAACACCTGTGGGCTAGCTAAAGCAATCTGA 1785

RESULT 14
ID AAN30049 standard; cDNA; 1314 BP.
AC AAN30049;
XX
XX 14-JUN-1992 (first entry)
DT
DE cDNA sequence corresponding to one of the allelic forms (B) of
XX bovine preprochymosin.
XX Chymosin; enzyme; rennet; cheese; ss.
XX
XX Bos taurus.
OS
FH Key Location/Qualifiers
FT sig_peptide 24..40
FT mat_peptide /*tag= a
FT 41..1169
FT /*tag= b
XX
XX BP77109-A.
XX
XX 20-APR-1983.
XX
XX 13-OCT-1982; 82BP-0201272.
XX
XX 14-OCT-1981; 81GB-0031004.
XX
XX (UNIL ) UNILEVER NV.
XX
XX Maat J, Verrips CT, Ledebor AM, Edens L;
XX
XX WPI; 1983-39656K/17.
XX
XX P-PSDB; AAP30086.
XX
XX DNA molecules comprising genes for preprochymosin - used to
XX transform microorganisms to give strain producing the
XX prepro-enzyme and its allelic and maturation forms
XX
XX Claim 2; Fig 1; 53bp; English.
XX
XX Preprochymosin is an intermediate (via prochymosin and
XX pseudochymosin) for the enzyme chymosin, which is the essential milk-
XX clotting component of rennet and is used in cheese manufacture.
XX AAN30049 corresp. to mRNA isolated from the fourth stomach of a
XX preumnant calf (abomasum, Frisian cow).
XX
XX Sequence 1314 BP; 309 A; 398 C; 338 G; 269 T; 0 other;
SO
Query Match 75.9%; Score 890.8; DB 4; Length 1314;
Best Local Similarity 87.3%; Pred. No. 4e-271;
Matches 976; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 56 TCCTGTCTTACTACAGCTGCTGAGATCAACCCGCAATTCCTCTCTCAAGTAAGTACTC 115
```

```
DB 52 TCTTGCCTCTCTCCAAAGCGCTGAGATTCACAGAGATCCCTCTGTATCAAAAGCAAGTCTC 111
QY 116 TCCGTAAAGCGCTGAAGAAATGATCTTGTAGAAATCTTTGCAAAATCAACAGTATG 175
DB 112 TGAGAAAGCGCTGAAGGAGATGCGCTTCTGAGAGATCTTCTGCAAAACAGCAGTATG 171
QY 176 GCATCGACGCAAGTACTCCGCTTCGAGAAAGTTGTAGCGGCGCACTTCAACTAC 235
DB 172 GCATCGACGCAAGTACTCCGCTTCGAGAAAGTTGTAGCGGCGCACTTCAACTAC 231
QY 236 TTGATAGTCAATATCTTGGGAAGATCTACCTCGAACCCTGCTCAAGATTCAACCGTTC 295
DB 232 TGATAGTCAAGTACTTGGGAAGATCTACCTGGGACCCCGCTCAAGATTCAACCGTTC 291
QY 296 TCTTGTATATCTGCTTCTGACTTCTGGGTTCCCTCTATCTATCTGCAAGAGATGCT 355
DB 292 TGTTTGACACTGCTCTCTGACTTCTGGGTTCCCTCTATCTATCTGCAAGAGCAATGCT 351
QY 356 GCAAGAACCAACCAAGATTCGATCCGAGAAAGTCGTCACCTTCCAGAACTTAAAGCAAC 415
DB 352 GCAAAACCAACCAAGCTTCGACCCGAGAAAGTCGTCACCTTCCAGAACTGAGCAAC 411
QY 416 CTTTGTCTATACACTACGCTACAGTACGATGCAAGAAATCTTAAAGCTATGATCCGCTCA 475
DB 412 CCCTGTCTATCCACTACGAGGACAGGACGATGCAAGGCAATCTGGGCTATGACCCGCTCA 471
QY 476 CTGTCTCAACATTTGTGACATTTCAACAGACATGAGACTTACGACCAAGAACAGGTG 535
DB 472 CTGTCTCAACATTTGTGACATTTCAACAGACATGAGGCTTGAGACCCAGAGACCCGGGG 531
QY 536 ATGTCTTCAACATGAGAAATTCAGATGCAATCTTGTATGCAATTCATGCTGCTGCT 595
DB 532 AGCTCTTCAACATGAGAAATTCAGATGCAATCTTGTATGCAATTCATGCTGCTGCTGCT 591
QY 596 CAGAGTACTGATACCTGTGTTTGAACAATGATGAACCGACACTAGTAGTCAAGACT 655
DB 592 CACAGTACTGATACCTGTGTTTGAACAATGATGAACCGACACTAGTAGTCAAGACT 651
QY 656 TGTTCCTGTTTACATGACAGGAATGCTCAGAGAGCATGCTACGCTTGGAGCTATTG 715
DB 652 TGTTCCTGTTTACATGACAGGAATGCTCAGAGAGCATGCTACGCTTGGAGCTATTG 711
QY 716 ATTCATCTTACATCAACAGATCTCTTCACTGCTTCAATCACTGTCAGAGTATGCTG 775
DB 712 ACCGTCCTTACATCAACAGATCTCTTCACTGCTTCAATCACTGTCAGAGTATGCTG 771
QY 776 AATTCACTGAGCAGATCTCAACATCAGCGGTGTGTTGTGATATGTAAGGTGATGTC 835
DB 772 AGTTCACTGAGCAGATCTCAACATCAGCGGTGTGTTGTGATATGTAAGGTGATGTC 831
QY 836 AAGCTATCTTGTGATACCGGTACGTCCTCAAGCTGTCGACCTTACAGACGATTTCAACA 895
DB 832 AGGCACTCTGAGCAGCGGCACTCTCAAGCTGTCGACCTTACAGACGATTTCAACA 891
QY 896 TTGAGCAAGTATGAGGCAACAGAACCAAGTACGCTGATTTGACATGATTTGGAGA 955
DB 892 TCCAGAGGCAATTTGAGGCAACAGAACCAAGTACGCTGATTTGACATGATTTGGAGA 951
QY 956 ACCTAGCTACATGCTCAAGTGTCTTGAAGTCAACGCGCAAGATGTACCCACTGACC 1015
DB 952 ACTTGAGCTACATGCTCAAGTGTCTTGAAGTCAATGCGCAAAATGTACCCACTGACC 1011
QY 1016 CTTCCGCTATACAGCAGAGATCAAGGGTCTGCACCAAGTGTATTCAGATGAGAAC 1075
DB 1012 CTTCCGCTATACAGCAGAGATCAAGGGCTTCTGTATCAAGTGTCTTCAGAGTGAATTC 1071
QY 1076 ATTCCCAAGAAATGATCTTTGGAGATGTGTTCAATCTGTAGTACTACAGCGCTTTGACA 1135
DB 1072 ATTCCCAAGAAATGATCTTTGGAGATGTGTTCAATCTGTAGTACTACAGCGCTTTGACA 1131
QY 1136 GGCGCAACACCTGTGGGCTAGCTAAAGCAATCTGA 1173
```

ID	Accession	Standard	DNA	Length
Db	1132	GGGCCAACACCTGCTGGGCTGGCCAAAGCATCTGA	1169	
XX	RESULT 15			
XX	AAN30022	AAN30022 standard; DNA; 1289 BP.		
XX	AAN30022;			
DT	25-MAR-2003	(updated)		
DT	25-APR-1992	(first entry)		
XX	Sequence of veal chymosin gene.			
DE	Proteolytic enzyme; zymogen; rennin; chymosin; cheese making;			
KM	microbial vector; ss.			
XX	Bos taurus.			
XX	Key	Location/Qualifiers		
FT	CDS	1..1140		
FT	/*tag=	a		
XX	BE897201-A.			
XX	03-NOV-1983.			
XX	30-JUN-1983;	83BE-0017731.		
XX	13-APR-1983;	83US-0484539.		
XX	01-JUL-1982;	82US-0394433.		
XX	(GENEX) GENEX CORP.			
DR	WPI; 1983-820813/47.			
DR	P-PSDB; AAP30013.			
XX	Isolated chymosin or rennin and prochymosin genes - plasmid(s)			
PT	which replicate in prokaryotic organisms, esp. Escherichia coli,			
PT	and organisms used for chymosin biosynthesis			
XX	Claim 7; Page 33-36; 43pp; French.			
XX	The inventors claim isolated chymosin (rennin) and prochymosin genes			
CC	from calves, and plasmids contg. the genes which are capable of			
CC	replicating in a prokaryotic organism. The prokaryotic organism is			
CC	pref. an Escherichia species, esp. E. coli p Gx 1225 (NRRL B-15061).			
CC	The microorganisms transformed by the plasmid are also claimed.			
CC	(Updated on 25-MAR-2003 to correct PA field.)			
XX	Sequence 1289 BP; 310 A; 391 C; 327 G; 260 T; 1 other;			
XX	Query Match	75.9%; Score 889.8; DB 4; Length 1289;		
XX	Best Local Similarity	87.2%; Pred. No. 8.2e-271;		
XX	Matches 975; Conservative	0; Mismatches 143; Indels 0; Gaps 0		
QY	56	TCGTTGCTGTTACTACAGCGCTGTAGATACACCGCATTCCTCTACAAAGTAAGTCTC	115	
Db	23	TCTTCGCTCTCTCCAGGCGCTGTAGATACCAAGATCCCTCTGACAAAGCAAGTCTC	82	
QY	116	TCCGTAAAGCGCTGAAAGGAACATGACTTCTAGAAAGCTTCTTGCAAAACACAGATAG	175	
Db	83	TGAGGAAGCGCTGAAAGGAGCATGGGCTTCTGGAAGACTTCTTGCAAAACAGAGTCTG	142	
QY	176	GCATGAGAGCAAGATACCTCCGCTTCGGTGAAGTGTCTAGAGTGTGCACCTTACCACTAC	235	
Db	143	GCATGAGAGCAAGATACCTCCGCTTCGGGAGTGTGCAGAGTGTCCCTGACCAACTAC	202	
QY	236	TTGATAGTCAATACTTTGGGAAAGATCTACCTCGAAACCCGCGCTCAAGATTCACCGTTC	295	
Db	203	TAGATAGTCAAGTCTTTGGGAAAGATCTACCTCGGAGCCCGGCCAGAGATTCACCGTTC	262	
QY	236	TCTTTGATACTAGTCTCTCTGACTTCTGGGTTCCCTCTATCTACTGCAAGAGCAATGCTC	355	

Db	263	TGTTGACACTGGCTCTCTGACTTCTGGGTACCTCTAATCTACTGCAAGACAATGCTT	322
Oy	356	GCAAGAACCAACCAAGATTGATCCGAGAAAGTCGTCACTTCCAGAACTTAGGCAAC	415
Db	323	GCAAAAACCAACGCGCTTGACCCCGAAGAAAGTCGTCACTTCCAGAACTTGGGCAAGC	382
Oy	416	CCCTGTATATACATAACGGTACAGGTAGTCATGCAAGGAATCTTAAGGTATGATACGGTCA	475
Db	383	CCCTGTATATACATAAGGAGACAGGCAGATGCAAGGCATCTTGGGCTTAGACACGGTCA	442
Oy	476	CTGTCTCCAACTATGTGACATTCAACAGACAGTAGACCTTAGACCCCAAGAACAGAGTG	535
Db	443	CTGTCTCCAACTATGTGACATCCAGACAGACAGTAGAGCTGAGAGACCAAGAGCCCGGGG	502
Oy	536	ATGTCTTACACTATGCAGAATTGATGACATCTTGGTATGCAATACCATCGTCCGCT	595
Db	503	ACGCTTTCACCTATGCGAAATTGCACGGGATCTCTGGGGATAGGCTTACCCCTCGCTCCCT	562
Oy	596	CAGAGTACCTGATACCTGTGTTTGAACAATGATGAAACGCAACACTAGTAGTCAACACT	655
Db	563	CAGAGTACCTGATACCTGTGTTTGAACAATGATGAAACGCAACCTGTGTGCCCAACACC	622
Oy	656	TGTTCTCGTTTTACATGACAGGAATGGCCAGAGAGACATGCTCACGCTTGAGACTATTG	715
Db	623	TGTTCTCGTTTTACATGACAGGAATGGCCAGAGAGACATGCTCACGCTTGGGGCCATCG	682
Oy	716	ATCCATCTTACTACACAGATCTCTTTCACTGGGTTCCAGTCACTGTGCACAGACTGCGC	775
Db	683	ACCGTCTCTACTACACAGGGTCTCCTGCACTGGGTGCCCGTGACAGTGCACAGACTGCGC	742
Oy	776	AATTCACTGAGCAGGTACCAATCAACGAGCTGTGTGTGTGATGTGAAAGTGGAGATGTC	835
Db	743	AGTTCACCTGAGCAGGTACCAATCAACGAGCTGTGTGTGTGATGTGAAAGTGGAGATGTC	802
Oy	836	AAAGTATCTTGATATACCGGTACGTCGCAAGCTGTGCGACCTTAGACAGCGAATTTCTACAA	895
Db	803	AGGCAATCTTGATATACCGGGACCTCCAAAGTTGGTGGGGCCCAAGAGGACATCTCTACAA	862
Oy	896	TTACAGCAAGCTATTGAGCAACACAGAACCAAGTACGGTGAATTGACATAGATTGGACA	955
Db	863	TCCAGCAGGGCAATTGGAGCCACACAGAACCAAGTACGGTGAATTGACATAGACTGCCACA	922
Oy	956	ACCTTACTCTACATGCCACAGTTGTCTTTAGATCAACGGCAAGATGTATACCACTGACCC	1015
Db	923	ACCTTACTCTACATGCCACACTGTGTCTTTGAGATCAATGGCAAAATGTATACCACTGACCC	982
Oy	1016	CTTCCGCTTATACACGACAGAGATCAAGGTTCTGCAACAGATGATTTCCAGAGTGGAAAC	1075
Db	983	CTTCCGCTTATACACGACCAAGACCAAGGGCTTCTGTACCAATGGCTTCCAGAGGAAATTC	1042
Oy	1076	ATTCCCAAGAAATGATCTTTGGAGATGTGTTCAATTGCTGAGTACTACAGGCTCTTTGACA	1135
Db	1043	ATTCCCAAGAAATGATCTTGGGGATGTTTTCATCCAGAGATATTACAGGCTCTTTGACA	1102
Oy	1136	GGGCAACAACCTGTGGGCTAGCTTAACCAATCTGA	1173
Db	1103	GGGCAACAACCTGTGGGCTGGCCCAACCAATCTGA	1140

Search completed: February 5, 2004, 03:13:10
Job time : 398 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2004, 02:57:45 ; Search time 92 Seconds
(without alignments)
5627.635 Million cell updates/sec

Title: US-09-643-755B-1
Perfect score: 1173
Sequence: 1 atgaactccttaagctctt.....ggtacgaacgaatcga 1173

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	897.2	76.5	1240	US-08-240-372-2	Sequence 2, Appl1
2	894.8	76.3	2732	US-08-846-021A-6	Sequence 6, Appl1
3	882	75.2	2732	5217891-14	Patent No. 5217891
4	203.6	17.4	832	US-08-318-193-15	Sequence 15, Appl1
5	203	17.3	828	5200327-6	Patent No. 5200327
6	199.8	17.0	846	5200327-5	Patent No. 5200327
7	199.4	17.0	838	US-08-318-193-13	Sequence 13, Appl1
8	192.2	16.4	2073	US-09-032-523-6	Sequence 6, Appl1
9	146.6	12.5	1299	US-08-723-938-4	Sequence 4, Appl1
10	146.6	12.5	1299	US-09-080-538-4	Sequence 4, Appl1
11	146.6	12.5	1299	US-09-387-413-4	Sequence 4, Appl1
12	146.6	12.5	1910	US-08-974-691-7	Sequence 7, Appl1
13	146.6	12.5	2061	US-09-008-271A-16	Sequence 16, Appl1
14	146.6	11.8	2038	US-08-631-097-7	Sequence 7, Appl1
15	138.2	11.8	2038	US-08-810-712-11	Sequence 11, Appl1
16	138.2	11.8	2465	PCT-US92-08080-1	Sequence 1, Appl1
17	128.6	11.0	1329	US-09-705-448-4	Sequence 4, Appl1
18	128.6	11.0	1329	US-09-705-448-4	Sequence 4, Appl1
19	128.6	11.0	1353	US-08-974-691-1	Sequence 1, Appl1
20	128.2	10.9	600	US-09-669-751-142	Sequence 142, Appl1
21	110	9.4	1615	US-08-640-305-5	Sequence 5, Appl1
22	110	9.4	1615	US-08-360-673-5	Sequence 5, Appl1
23	109	9.3	2875	US-08-328-314-1	Sequence 1, Appl1
24	109	9.3	2875	US-08-731-045-1	Sequence 1, Appl1
25	97.8	8.3	2032	US-08-088-633-1	Sequence 1, Appl1
26	97.8	8.3	2032	US-08-245-756-1	Sequence 1, Appl1
27	97.8	8.3	2032	US-08-441-750-1	Sequence 1, Appl1

28	97.8	8.3	2032	US-08-441-751-1	Sequence 1, Appl1
29	97.8	8.3	2032	PCT-US92-02521-1	Sequence 1, Appl1
30	95.8	8.2	2454	US-09-079-415-1	Sequence 1, Appl1
31	74.8	6.4	392	US-09-702-705-1157	Sequence 1157, Appl1
32	74.8	6.4	392	US-09-736-457-1157	Sequence 1157, Appl1
33	67.6	5.8	663	US-09-280-116-265	Sequence 265, Appl1
34	63.2	5.4	227	US-09-705-448-6	Sequence 6, Appl1
35	57.6	4.9	264	US-09-016-434-292	Sequence 292, Appl1
36	52.6	4.5	80	US-08-146-422-23	Sequence 23, Appl1
37	52.6	4.5	80	US-08-146-422-24	Sequence 24, Appl1
38	52.6	4.5	80	US-08-626-554-5	Sequence 5, Appl1
39	52.6	4.5	80	US-08-626-554-29	Sequence 29, Appl1
40	52.6	4.5	80	US-08-693-709-6	Sequence 6, Appl1
41	51.6	4.4	80	US-09-230-590-4	Sequence 4, Appl1
42	51.4	4.4	900	US-08-181-271A-4	Sequence 4, Appl1
43	51.4	4.4	900	US-08-449-315-4	Sequence 4, Appl1
44	51.4	4.4	900	US-08-444-803-4	Sequence 4, Appl1
45	51.4	4.4	900	US-08-449-043-4	Sequence 4, Appl1

ALIGNMENTS

```
RESULT 1
US-08-240-372-2
; Sequence 2, Application us/08240372
; Patent No. 5741665
; GENERAL INFORMATION:
; APPLICANT: KATO, ELIE K.
; APPLICANT: STUART, W. DORSEY
; TITLE OF INVENTION: LIGHT-REGULATED PROMOTERS FOR PRODUCTION
; TITLE OF INVENTION: OF HETEROLOGOUS PROTEINS IN FILAMENTOUS FUNGI
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSER: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,372
; FILING DATE: 10-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 3918-0003.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; US-08-240-372-2
;
; Query Match 76.5%; Score 897.2; DB 1; Length 1240;
; Best Local Similarity 87.7%; Pred. No. 8.9e-310;
; Matches 980; Conservative 0; Mismatches 138; Indels 0; Gaps 0;
;
; QY TCGTTCCTGCTACTACGCTGTGAGTACCCGATCTCTCTACAAAGTATGCTC 115
; DB TCTTGCTCTCTCCAGGCGCTGAGTACCGATCTCTGTACAAAGGAGTCTC 159
; QY 116 TCCGTAAAGCCCTGAAGGAGATGACTTCTAGAGACTTCTTGAGAAACAGATAG 175
```

```

Db      160 TGAGGAAGGGCTGGAAGAGCATGGGCTTTGAGAGGACTTCTGACAGAAACAGCATATG 219
QY      176 GCATGAGCAAGAGTACTCCGGCTTGGTGAAGTTGCTAGCGTGCACCTTACCAACTACC 235
Db      220 GCATGAGCAAGAGTACTCCGGCTTGGGAGAGGTGGCAGCGTGCCCTGACCAACTACC 279
QY      236 TTGATAGTCAATCTTTGGGAAGATCTACTCGGAAACCCCGCTCAAGAGTTCAACGGTTC 295
Db      280 TGGATAGTCAATCTTTGGGAAGATCTACTCGGAAACCCCGCTCAAGAGTTCAACGGTTC 339
QY      296 TCTTTGATAGTGGTTCCTCTGACCTTCTGGGTTCCCTCTATCTATCTGCAAGAGCAATGGCT 355
Db      340 TGTTCACACTGGCTCTCTGACTTCTGGGTTACCTCTTATCTATCTGCAAGAGCAATGGCT 399
QY      356 GCAAGAACCAACCAAGATTGATCCGAGAAAGTGTCCACTTTCAGAACTTAGAGCAAAAC 415
Db      400 GCAAAACCAACCAAGCGCTTGACCCGAGAAAGTGTCCACTTTCAGAACTTAGAGCAAAAC 459
QY      416 CCTTGTCTATACCTACGTTACGTTACGTTACGTTACGTTACGTTACGTTACGTTACGTTAC 475
Db      460 CCTTGTCTATACCTACGTTACGTTACGTTACGTTACGTTACGTTACGTTACGTTACGTTAC 519
QY      476 CTGTCTCCAACTTGTGACATTCACAGACAGTAGGACTTAGAACCCCAAGAACAGGTG 535
Db      520 CTGTCTCCAACTTGTGACATTCACAGACAGTAGGACTTAGAACCCCAAGAACAGGTG 579
QY      536 ATGTCTTCACTTATGAGAAATTCAGTGCATCTTGTGTATGGCAATCCATCGCTCGCT 595
Db      580 AGTCTTCACTTATGAGAAATTCAGTGCATCTTGTGTATGGCAATCCATCGCTCGCT 639
QY      596 CAGAGTACTGATCTGTGTGTGTTGACAACTGATGAAACCACTAGTACTGAACT 655
Db      640 CAGAGTACTGATCTGTGTGTGTTGACAACTGATGAAACCACTAGTACTGAACT 699
QY      656 TGTTCCTGGTTTATGAGACAGGAATGGCAGAGAGCATGCTAGCGTGTGAGTATTG 715
Db      700 TGTTCCTGGTTTATGAGACAGGAATGGCAGAGAGCATGCTAGCGTGTGAGTATTG 759
QY      716 ATCCATCTCTATACAGAGATCTCTTCACTAGGTTCCAGTCACTGTGACAGCACTAGTGC 775
Db      760 ACCGCTCTATACAGAGATCTCTTCACTAGGTTCCAGTCACTGTGACAGCACTAGTGC 819
QY      776 AATTCACTGTGACAGTGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 835
Db      820 AGTTCACTGTGACAGTGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 879
QY      836 AAGCTATCTTGAATCCGGTACGTTCAAGCTGTGCGACCTTAGCAGAGCAATCTCAACA 895
Db      880 AAGCTATCTTGAATCCGGTACGTTCAAGCTGTGCGACCTTAGCAGAGCAATCTCAACA 939
QY      896 TTCAGCAAGCTATGAGACCAAGAACAGTACGTTGATTTGACATAGTTGACACA 955
Db      940 TTCAGCAAGCTATGAGACCAAGAACAGTACGTTGATTTGACATAGTTGACACA 999
QY      956 ACCTTGAAGTACGTTGACAGTGTCTTTGAGATCAACGGCAAGTATCCACTGACCC 1015
Db      1000 ACCTTGAAGTACGTTGACAGTGTCTTTGAGATCAACGGCAAGTATCCACTGACCC 1059
QY      1016 CCTCCGCTATACAGCAGAGATCAAGGGTTCTGACCACTGATTCACAGTGAAGAAC 1075
Db      1060 CCTCCGCTATACAGCAGAGATCAAGGGTTCTGATCAAGTGTCTTCAAGTGAAGAAC 1119
QY      1076 ATTCGCAAGAAATGATCTTGGGAGATGTTCATTCGTGAGTACTACAGCGTCTTTGACA 1135
Db      1120 ATTCGCAAGAAATGATCTTGGGAGATGTTCATTCGTGAGTACTACAGCGTCTTTGACA 1179
QY      1136 GGGCCAAACAACCTGTTGGGCTAGCTAAAGCAATCTGA 1173
Db      1180 GGGCCAAACAACCTGTTGGGCTAGCTAAAGCAATCTGA 1217

```

RESULT 2

```

US-08-846-021A-6
; Sequence 6, Application US/08846021A
; Patent No. 5948682
; GENERAL INFORMATION:
; APPLICANT: Moloney, Maurice M.
; TITLE OF INVENTION: Preparation of Heterologous Proteins on
; TITLE OF INVENTION: Oil Bodies
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; City: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/846,021A
; FILING DATE: April 25, 1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Graville, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 9369-039
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 850..1206
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1444..2729
; US-08-846-021A-6

Query Match 76.3%; Score 894.8; DB 2; Length 2733;
Best Local Similarity 88.4%; Pred. No. 1.2e-308;
Matches 971; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY      76 GCTGAGATCAACCGCATTCCTCTCTTACAAAGTAACTCTCCGTAAGGCGCTGAAGAA 135
Db      1630 GCTGAGATCAACCGCATTCCTCTCTTACAAAGGCAAGTCTCTGAGAAAGGCGTGAAGAG 1689
QY      136 CATGAACTTCTAGAAAGCTTTCTTGCAAAACAGATATGCACTGACGAGCAAGTACTCC 195
Db      1690 CATGAACTTCTAGAAAGCTTTCTTGCAAAACAGATATGCACTGACGAGCAAGTACTCC 1749
QY      196 GGGCTGGGGAAGTGTCTAGCGCGTGCACCTTACCACTACTGATGTGTAATCTTTGGG 255
Db      1750 GGGCTGGGGAAGTGTCTAGCGCGTGCACCTTACCACTACTGATGTGTAATCTTTGGG 1809
QY      256 AAGATCTACTCGGAACCCCGCTCAAGAGTTCAACGGTCTCTTTGATAGTGGTCTCT 315
Db      1810 AAGATCTACTCGGAACCCCGCTCAAGAGTTCAACGGTCTCTTTGATAGTGGTCTCTCT 1869
QY      316 GACTTGTGGGTTCTCTTATCTTACGCAAGAGCAATGCTGCAAGAACCAACCAAGATTTC 375
Db      1870 GACTTGTGGGTTCTCTTATCTTACGCAAGAGCAATGCTGCAAGAACCAACCAAGATTTC 1929
QY      376 GATCCGAAAGAGTGTCTCACTTCCAGAACTTAGGCAAAACCTTGTCTATACCTAGGT 435
Db      1930 GATCCGAAAGAGTGTCTCACTTCCAGAACTTAGGCAAAACCTTGTCTATACCTAGGT 1989

```


Db 1689 ATTCCAGAAATGATCTGGGGAGTGTTCATCCGAGATATACAGCGCTTTGACA 1748

Qy 1136 GGGCAACAACCTCGTTGGCTAGCTAAGCAATCTGA 1173

Db 1749 GGGCAACAACCTCGTTGGCTAGCTAAGCAATCTGA 1786

RESULT 4

US-08-318-193-15

Sequence 15, Application US/08318193

Patent No. 5641663

GENERAL INFORMATION:

APPLICANT: GARVIN, Robert T.

APPLICANT: MALEK, Lawrence T.

TITLE OF INVENTION: AN EXPRESSION SYSTEM FOR THE SECRETION

TITLE OF INVENTION: OF BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY

TITLE OF INVENTION: STIMULATING FACTOR (GM-CSF) AND OTHER HETEROLOGOUS

TITLE OF INVENTION: PROTEINS FROM STREPTOMYCES

NUMBER OF SEQUENCES: 91

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: Virginia

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/318,193

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,314

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 18740/116 CACO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 832 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: Other nucleic acid;

DESCRIPTION: Synthetic DNA oligonucleotide

FEATURE:

NAME/KEY: CDS

LOCATION: 399..830

US-08-318-193-15

Query Match 17.4%; Score 203.6; DB 1; Length 832;

Best Local Similarity 75.1%; Pred. No. 4.5e-62;

Matches 254; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy 57 CGTTGCTGTACTACAGCTGTGATGACACCCGATTCCTCTACAAAGTAACTCT 116

Db 494 CGTTCCACCGCAACGCTGCAAGATCACTAGATCCCATTTGTACAGGTAAGTCTT 553

Qy 117 CGTAAAGCGCTGAAGAACATGACTTTCTAGAACTTTCTTGAGAAACAAGTATG 176

Db 554 GAGAAAGCGCTTGAAGAACACGTTTGTGGAAGACTTTCTTGAGAAACAAGTATG 613

Qy 177 CATCAGCAGCAAGTACTCCGGCTTCGCTGAAGTTGCTAGCGCACTTACCAATACCT 236

Db 614 TATCTCTCAAGTACTGTGTTGGTAAGTCGTTCCGTTCCATTACCAACTACTT 673

Qy 237 TGAATGTAATACCTTTGGAGATCTACCTCGAAACCCGCTCAAGATTCACCGTCT 296

Db 674 GACTCCCAATACCTTCGTAAGATCTACTAGTACCCACCAAGAAATTCAGTCTT 733

Qy 297 CTTGATCTAGTCTCTGACTTCTGAGTCTGGGTTCCCTATCTACTGCAAGACATCGCTG 356

Db 734 GTTCGACACCGGTTCTTCTGACTTCGGGTCCATCGATGATTTACTGTAAGTCCACGCTTG 793

Qy 357 CAAGAACCAACAAGTTGATCCGAAAGTCGCA 394

Db 794 TAAGAACCAACAAGTTGATCCGAAAGTCTCTA 831

RESULT 5

5200327-6

Patent No. 5200327

APPLICANT: GARVIN, ROBERT T.; MALEK, LAWRENCE T.

TITLE OF INVENTION: EXPRESSION SYSTEM FOR THE SECRETION OF

BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY STIMULATING

FACTOR (GM-CSF) AND OTHER HETEROLOGOUS PROTEINS FROM

STREPTOMYCES

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/224,568

FILING DATE: 26-JUL-1988

SEQ ID NO: 6:

LENGTH: 828

5200327-6

Query Match 17.3%; Score 203; DB 6; Length 828;

Best Local Similarity 75.8%; Pred. No. 7.3e-62;

Matches 251; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 57 CGTTGCTGTACTACAGCTGTGATGACACCCGATTCCTCTCTACAAAGTAACTCT 116

Db 494 CGTTCCACCGCAACGCTGCAAGATCACTAGATCCCATTTGTACAGGTAAGTCTT 553

Qy 117 CGTAAAGCGCTGAAGAACATGACTTTCTAGAACTTTCTTGAGAAACAAGTATG 176

Db 554 GAGAAAGCGCTTGAAGAACACGTTTGTGGAAGACTTTCTTGAGAAACAAGTATG 613

Qy 177 CATCAGCAGCAAGTACTCCGGCTTCGCTGAAGTTGCTAGCGCACTTACCAATACCT 236

Db 614 TATCTCTCAAGTACTGTGTTGGTAAGTCGTTCCGTTCCATTACCAACTACTT 673

Qy 237 TGAATGTAATACCTTTGGAGATCTACCTCGAAACCCGCTCAAGATTCACCGTCT 296

Db 674 GACTCCCAATACCTTCGTAAGATCTACTAGTACCCACCAAGAAATTCAGTCTT 733

Qy 297 CTTGATCTAGTCTCTGACTTCTGAGTCTGGGTTCCCTATCTACTGCAAGACATCGCTG 356

Db 734 GTTCGACACCGGTTCTTCTGACTTCGGGTCCATCGATGATTTACTGTAAGTCCACGCTTG 793

Qy 357 CAAGAACCAACAAGTTGATCCGAAAGTCGCA 387

Db 794 TAAGAACCAACAAGTTGATCCGAAAGTCTCTA 824

RESULT 6

5200327-5

Patent No. 5200327

APPLICANT: GARVIN, ROBERT T.; MALEK, LAWRENCE T.

TITLE OF INVENTION: EXPRESSION SYSTEM FOR THE SECRETION OF

BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY STIMULATING

FACTOR (GM-CSF) AND OTHER HETEROLOGOUS PROTEINS FROM

STREPTOMYCES

NUMBER OF SEQUENCES: 24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/224,568

FILING DATE: 26-JUL-1988

SEQ ID NO:5:
LENGTH: 846
5200327-5

Query Match 17.0%; Score 199.8; DB 6; Length 846;
Best Local Similarity 76.2%; Pred. No. 1e-60;
Matches 246; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 74 CTGCTGAGATGACCCGCAATTCCTCTCTCAAGGTAAGCTCTCCGTAGAGGGCTGAAG 133
DB 517 CTGCGAGATGACTGTAATCCATTGTACAGGGTAAGCTTTGAGAAAGGCTTGAAGG 576
QY 134 AACATGACTTCTAGAGACTTCTTGCAAGAAACAAGTATGCAATCAGCAGCAACTACT 193
DB 577 AACACGGTTTGTGAGAGACTTCTTGCAAAAGCAACAATAGGTAATCTCTCCAGTACT 636
QY 194 CCGGCTTGCGTGAAGTTGCTAGCGTGCACCTTACCACTACTGTAATGTAATCTTGG 253
DB 637 CTGGTTGCGTGAAGTCCGCTCCGTTCCATTGACCAACTACTTGGACTCCCAATCTTCG 696
QY 254 GGAAGATCTACTCGGAACCCGCGCTCAAGAGTTGACCGCTTCTTGAATAGTCTTCT 313
DB 697 GTAAGATCTACTAGTACCCCAACAAGAAATTCATGCTTGTTCGACACCGGTTCTT 756
QY 314 CTGACTTCTGGGTTCCCTCTATCTACTGCAAGAGCAATGCTGCAAGAACCAACAAGAT 373
DB 757 CTGACTTCTGGGTTCCCATGATTTACTGTAAGTCCAGCGTTGTGAAGAACCAACAAGAT 816
QY 374 TCGATCCGAGAAAGTCGTCACC 396
DB 817 TCGACCAAGAAAGACTCTAGC 839

RESULT 7
US-08-318-193-13
Sequence 13, Application US/08318193
Patent No. 5641663

GENERAL INFORMATION:

APPLICANT: GARVIN, Robert T.

APPLICANT: MALEK, Lawrence T.

TITLE OF INVENTION: AN EXPRESSION SYSTEM FOR THE SECRETION

TITLE OF INVENTION: OF BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY

TITLE OF INVENTION: STIMULATING FACTOR (GM-CSF) AND OTHER HETEROLOGOUS

NUMBER OF SEQUENCES: 91

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: Virginia

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/318,193

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,314

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 18740/116 CACO

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-9300

TELEFAX: (703) 683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 838 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: Other nucleic acid;

DESCRIPTION: Synthetic DNA oligonucleotide

FEATURE:

NAME/KEY: CDS

LOCATION: 399..836

US-08-318-193-13

Query Match 17.0%; Score 199.4; DB 1; Length 838;

Best Local Similarity 76.3%; Pred. No. 1.4e-60;

Matches 245; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 74 CTGCTGAGATGACCCGCAATTCCTCTCTCAAGGTAAGCTCTCCGTAGAGGGCTGAAG 133
DB 517 CTGCGAGATGACTGTAATCCATTGTACAGGGTAAGCTTTGAGAAAGGCTTGAAGG 576
QY 134 AACATGACTTCTAGAGACTTCTTGCAAGAAACAAGTATGCAATCAGCAGCAACTACT 193
DB 577 AACACGGTTTGTGAGAGACTTCTTGCAAAAGCAACAATAGGTAATCTCTCCAGTACT 636
QY 194 CCGGCTTGCGTGAAGTTGCTAGCGTGCACCTTACCACTACTGTAATGTAATCTTGG 253
DB 637 CTGGTTGCGTGAAGTCCGCTCCGTTCCATTGACCAACTACTTGGACTCCCAATCTTCG 696
QY 254 GGAAGATCTACTCGGAACCCGCGCTCAAGAGTTGACCGCTTCTTGAATAGTCTTCT 313
DB 697 GTAAGATCTACTAGTACCCCAACAAGAAATTCATGCTTGTTCGACACCGGTTCTT 756
QY 314 CTGACTTCTGGGTTCCCTCTATCTACTGCAAGAGCAATGCTGCAAGAACCAACAAGAT 373
DB 757 CTGACTTCTGGGTTCCCATGATTTACTGTAAGTCCAGCGTTGTGAAGAACCAACAAGAT 816
QY 374 TCGATCCGAGAAAGTCGTCAC 394
DB 817 TCGACCAAGAAAGACTCTA 837

RESULT 8
US-09-032-523-6
Sequence 6, Application US/09032523
Patent No. 6232454

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Guejler, Karl

APPLICANT: Baugh, Mariah

TITLE OF INVENTION: HUMAN PROTEINASE MOLECULES

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/032,523

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0479 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2073 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PANTU01
CLONE: 1515165
US-09-032-523-6

Query Match 16.4%; Score 192.2; DB 3; Length 2073;
Best Local Similarity 56.5%; Pred. No. 1e-57;
Matches 406; Conservative 0; Mismatches 298; Indels 15; Gaps 2;

```
QY 87 CCGCATTCCTCTACAAAGTAAGTCTCCGTAAGGCGCTGAAGAGATGACTTCT 146
DB 104 CAGGTCGCCCTCAGAGGCACTCGTCCCTCAGAGAGAGCTGGGCGACGAGCCAGCT 163
QY 147 AGAAGACTTCTTCAGAAACAAAGATGAGATCA-----GCAGCAAGTACTCCGCTT 200
DB 164 CTCGAGTTCTGGAATCCATATATTGACATGATCAGTTACCGAGTCTGCTCAAT 223
QY 201 CGGTGAAGTTCTGACGCTGCACTTACCACTACCTTGATGATGATGATGATGATGAT 260
DB 224 GAGCCAGAGGCAAGAGAACCCCTCATCACTACTTGATATGGAATATCTTCCGCTAT 283
QY 261 CTACCTCGAAGCCCGGCTCAAGAGTTCACCGTTCTTTGATATGAGTCTCTGACTT 320
DB 284 CTCATTTGCTCCCGACCAAGACTTTCATCTGTCATCTTGACACTGGCTCTCCAACT 343
QY 321 CTGGGTTCCCTCTATCTACTGCAAGAGCAATGCTGCAAGAACCAAGATTCATCC 380
DB 344 CTGGGTCCTCTGTCATGACACTGACCTGCAAGGCTGCAAGGCGACAGGTTCCAGCC 403
QY 381 GAGAAAGTCTCACTCTTCCAGAACTTAGGCAAACTTGTCTATATCACTACGATACG 440
DB 404 TTCCAGTCCAGACATACAGCCGACGAGTCAATCTTCTCATATGATGAGACCGG 463
QY 441 TAGCATGCAAGATCTTAGGCTATGATACCGTACGTCCTCCAACTTGTGACATTTCA 500
DB 464 GAGCTTGTCCGGATCATGAGCCGACCAAGTCTCTGTGAAAGACTAAACCGGTTGG 523
QY 501 ACAGACAGTAGACTTAGCAACCAAGAGATGATGCTTCACTATGCAAGATTCGA 560
DB 524 CCAGCAAGTTTGAAGAAAGTGTCAAGAGCCAGGCAAGACCTTTGTGATGCAAGTTTGA 583
QY 561 TGGCATCTTGTGATGCAATACCATGCTGCGCTCAAGATCTCGATACCTGTGTTGA 620
DB 584 TGGAAATCTGGGCTGGGATATCCCTCTTGTGCTGTGGAGAGATGACTCCAGATTTGA 643
QY 621 CAACATGATGAACCGACACTAGTAGTCAAGACTTGTCTCGGTTTCAATGACAGAA 680
DB 644 CAACATGATGATGATGAGACCTGGGAGCTTGGCGAGTCTTGTCTCAATGAGCAATGA 703
QY 681 -----TGGCAGAGAGCATGCTCAGCTTGTGAGCTATTGATCCATCTACTACAC 731
DB 704 CCAGAAAGTGTGTCGGGAGCGAGCTGATTTTGGAGGCTACAGACACTCCATTTCTC 763
QY 732 AGGATCTTCACTGAGTTCAGTCACTGTGACAGCACTACTGCAATTCACCTGTGACA 790
DB 764 TGGAGAGCTGAATTTGGGTTCCAGTCAACAGCAAGCTTACTGGGAGATGACACTGATA 822
```

RESULT 9

US-08-723-938-4
Sequence 4, Application US/08723938
Patent No. 5776759
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Coleman, Roger
TITLE OF INVENTION: TWO NOVEL HUMAN CATHESPIN PROTEINS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,938
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0125 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1299 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
IMMEDIATE SOURCE:
LIBRARY: LUNGN0102
CLONE: 312099
US-08-723-938-4

Query Match 12.5%; Score 146.6; DB 1; Length 1299;
Best Local Similarity 50.4%; Pred. No. 1.4e-41;
Matches 421; Conservative 0; Mismatches 399; Indels 15; Gaps 2;

```
QY 216 CGTGCACTTACCACTACTGATGATGCAATCTTTGGGAAAGATCTACTCGAAACCC 275
DB 217 CGTAACTCTCTGAACTACAGGAGATGCAATTTTGGGAAATTTGGGCTGGGAACGCC 276
QY 276 GCTTCAAGAGTTACCGTTCTCTTATATATGTTTCTGACTTCTGGGTTCCCTCTA 334
DB 277 TCCACAAACCTTACAGTGTGCTTTGACACTGGCTCTCCAACTCTGGGTCCTCCGTCAG 336
QY 335 -----TCTACTGCAAGCAATGCTGCAAGAACCAACCAAGATTCGATCCGAAAGTC 389
DB 337 GAGATGCCACTTTTCAAGTGTGCTGCTGCTGTTACACACAGATTTGATCCCAAAGCTTC 396
QY 390 GTCCACTTTCAGAACTTAGCAAAACCTTGTCTATATACATACGATACAGTACATGCA 449
DB 397 TACCTCTTTCAGGCAATGAGGACCAAGTTTGTGCAATATATGAACTGGGCGGGTAGA 456
QY 450 AGGAATCTTAGGCTATGATACGCTCACTGTCTTCCAACTTTGGAATTTCAACAGACAGT 509
```

```

Db      457 TGGAAATCTGAGCGAGGAGCAAGCTGATATTTGGTGGATCAAGGCTGATCATGATTTT 516
Qy      510 AGGACTTATGACCCCAAGCAAGTGTATGTTTCACTTATGAGATTCATGATGATCT 569
Db      517 CGGGGAGGCTCTCTGGAGGCCAGCTGTCTTCCGCTTTGGCCATTTTGAAGGGATATT 576
Qy      570 TGGTATGCGATACCCATCGCTCGCTCAGAGTACTGATACCTGTGTGTTGACAAATGAT 629
Db      577 GGGCCCTCGGTTTTTCCCATTTCTGTGTGGAAGAGTGGGCCCGGATGATGATCTGT 636
Qy      630 GAACCGACACTGATGACTCAAGACTTGTCTCGGTTTACATGACAGGAATGGCCAGGA 689
Db      637 GGAGCAGGGGCTATTGATGATAGCTGTCTCTCTTTTAACTTCAACAGGGAACCTGAAGA 696
Qy      690 G-----AGCATGCTCAGGCTTGGAGCTATGATTCATCTCTACTACAGGATCTCT 740
Db      697 GCTGTATGAGAGAGAGCTGTGCTCTGGGGGGCTCGGACCGGACACTATATCCACCCCT 756
Qy      741 TCACCTGGGTTCCAGTCACTGTGTCAGCAGTACTGCAATTCAGTGTGACAGTGCACCAT 800
Db      757 CACCTTGTGTCAGTCAAGGCTCCCGGCTACTGAGAGATCCATGAGAGCTGTGTAAGT 816
Qy      801 CAGCGGTGTGTTGTTGATGTAAGGTGATGTCAGTATCTTGGATACCGGTACGTC 860
Db      817 GGGCCCAAGGCTGACTCTCTGTGCAAGGGCTGTGCTGCATCTGGATACGGGACAGTC 876
Qy      861 CAAGCTGTGTCAGCTTAGCAGAGCACTTCTCAACATTCAGCAAGCTATTGGAGCCACACA 920
Db      877 CCTCATCACAGAACCCACTGAGAGATCCGGGCCCTGTCAGCAGCATTGGGGGAATCC 936
Qy      921 GAACCACTACGCTGATGATTGATGATGCGACAACCTTACATGCTTACAGCTTGT 980
Db      937 CTGTGCTGTGGGAGTATCATCTGTGCTGCGAAATCCAAAGCTCCCGGAGTCTC 996
Qy      991 CTTTGAATCAAGCGCAAGATGTACCACTGACCCCTCCGCTTATACAGGCGAG 1035
Db      997 CTTCCTTCTTGGGGGCTGTGTTTAACTTCAACGGCCCATGATACGTATCAG 1051

```

RESULT 10

```

; Sequence 4, Application US/09080538
; Patent No. 5965129
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Coleman, Roger
; TITLE OF INVENTION: TWO NOVEL HUMAN CATHESPIN PROTEINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/080,538
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/723,938
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0125 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555

```

```

; TELEFAX: 415-845-4166
; TELE:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1299 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; IMMEDIATE SOURCE:
; LIBRARY: LUNGMOT02
; CLONE: 312099
; US-09-080-538-4

```

```

Query Match      12.5%; Score 146.6; DB 2; Length 1299;
Best Local Similarity 50.4%; Pred. No. 1,4e-41;
Matches 421; Conservative 0; Mismatches 399; Indels 15; Gaps 2;

```

```

Qy      216 CGTGCACCTTACCACTTATGATGATCAATCTTTGGGAAGATCTTACCTGGAAACCC 275
Db      217 CGTACCTCTCTGAACTACAGGAGATGTCAGATATTTGGGAAATTTGGGCTGGAAACCC 276
Qy      276 GCTTCAAGATTCACCGTTCTCTTGTATGATCTGTCTCTGACTTCTGGGTTCCCTTA- 334
Db      277 TCCACAAATCTCATGTTGCTTTGACACTGGCTCTCCAAATCTGTGGTCCGTCAG 336
Qy      335 -----TCTACTGCAAGAGCAATGCTGCAAGAACCAACCAAGATTCATCCGAGAAATC 389
Db      337 GAGATGCCACTTCTTCACTGATGTCCTGCTGTGATACACCAAGATTTGATCCAAAGCCTC 396
Qy      390 GTCCACCTTCAGAACTTAGGCAAACTTGTCTATATACCTACGATACAGTATGATGCA 449
Db      397 TAGTCTCTTCAGGCAATGAGGACCAAGTTGGCATTCATATGGAATCGGGCGGTAGA 456
Qy      450 AGGATCTTAGGCTATGATACCGTACTGTCTCCAACTTGTGACATTCACAGACAGT 509
Db      457 TGGATCTTGAAGCGAGCAAGCTGACTATGTGGAATCAAGGTGATCATGATATT 516
Qy      510 AGGACTTACCAACCAAGAGTATGTCCTTCACTATGAGAAATTCATGATGATCT 569
Db      517 CGGGAGGCTCTCTGGAGCCAGCTGTGCTTGGCTTTTGGCCATTTGATGGATATT 576
Qy      570 TGGTATGCAATACCATGCTCGCTGAGATGATCTGATACCTGTGTTGACAAATGAT 629
Db      577 GGGCCTCGGTTTTCCCATTTCTGTGTGGAAGAGTTGGGCCCGGATGATGATCTGT 636
Qy      630 GAACCGACACTGATGACTCAAGACTTGTCTCGGTTTACATGACAGGAATGGCCAGGA 689
Db      637 GGAGCAGGGGCTATTGATGATAGCTGTCTCTCTTTTAACTTCAACAGGGAACCTGAAGA 696
Qy      690 G-----AGCATGCTCAGGCTTGGAGCTATGATTCATCTCTACACAGATCTCT 740
Db      697 GCTGTATGAGAGAGAGCTGTGCTCTGGGGGGCTCGGACCGGACACTATATCCACCCCT 756
Qy      741 TCACCTGGGTTCCAGTCACTGTGTCAGCAGTACTGCAATTCAGTGTGACAGTGCACCAT 800
Db      757 CACCTTGTGTCAGTCAAGGCTCCCGGCTACTGAGAGATCCATGAGAGCTGTGAAGT 816
Qy      801 CAGCGGTGTGTTGTTGATGTAAGGTGATGTCAGTATCTTGGATACCGGTACGTC 860
Db      817 GGGCCCAAGGCTGACTCTCTGTGCAAGGGCTGTGCTGCATCTCGATACGGGACAGTC 876
Qy      861 CAAGCTGTGTCAGCTTAGCAGAGCACTTCTCAACATTCAGAGAGTATTTGAGCCACACA 920
Db      877 CCTCATCACAGAACCCACTGAGAGATCCGGGCCCTGTATGACAGCATTGGGGGAATCC 936
Qy      921 GAACCACTACGCTGATGATTGATGATGCGACAACCTTACATGCTTACAGCTTGT 980
Db      937 CTGTGCTGTGGGAGTATCATCTGTGTCTCGGAATCCAAAGCTCCCGGAGTCTC 996

```


Sequence 7, Application US/08631097
Patent No. 5968816

GENERAL INFORMATION:

APPLICANT: Kimchi, Adi
TITLE OF INVENTION: Tumor Suppressor Genes,
TITLE OF INVENTION: Protein Encoded Thereby, and Use of Said Genes and Protein
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Migman, Cohen, Leitner, & Myers, P.C.
STREET: 900 17th Street, N.W., Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/631,097
FILING DATE: 12-Apr-96

CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/11598

FILING DATE: 12-Oct-94

ATTORNEY/AGENT INFORMATION:

NAME: Cohen, Herbert

REGISTRATION NUMBER: 25,109

REFERENCE/DOCKET NUMBER: 0744.057

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)463-7700

TELEFAX: (202)473-6915

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 2038 base pairs

TYPE: nucleic acid

STRANDEDNESS: Double

TOPOLOGY: Linear

MOLECULE TYPE: Genomic DNA

HYPOTHETICAL: No

ANTI-SENSE: No

FRAGMENT TYPE: No. 5968816 applicable

ORIGINAL SOURCE:

ORGANISM: homo sapiens

STRAIN: not applicable

INDIVIDUAL ISOLATE: not applicable

DEVELOPMENTAL STAGE: not applicable

HAPLOTYPE: not applicable

TISSUE TYPE: blood

CELL TYPE: Leucocyte

CELL LINE: HeLa

ORGANELLE: not applicable

IMMEDIATE SOURCE:

LIBRARY: not applicable

CLONE: not applicable

POSITION IN GENOME:

CHROMOSOME/SEGMENT: not applicable

MAP POSITION: not applicable

UNITS: not applicable

FEATURE:

NAME/KEY: This is the DNA sequence

LOCATION: claimed in 15(vi) as the Cathepsin gene in FIG. 15.

IDENTIFICATION METHOD: experiment-

IDENTIFICATION METHOD: in specification

OTHER INFORMATION: prevention of IFN-2

OTHER INFORMATION: promoted cell death

PUBLICATION INFORMATION: not available

US-08-631-097-7

Query Match 11.8%; Score 138.2; DB 2; Length 2038;
Best Local Similarity 51.0%; Pred. No. 2e-38;

Matches	443; Conservative	0; Mismatches	378; Indels	48; Gaps	3;
QY	216	CGTGCACCTTACCACTACCTTGAATGATACCTTGGGAAGATCTACCTCGAAGCC	275		
Db	255	CGAGTGTCTAAGAACTTACATGACGCGCAGTATCTACGGGAGATTTGGCATGCGGAC	314		
QY	276	GCTTCAGAGTTCACCGTTCTCTTTGATATGTTCTCTGATCTTGGGTTCCCTAT	335		
Db	315	CCCCAGTGTTCACAGTGTCTTTCAGACAGGGCTCTCCAACTGTGGTCCCTCCAT	374		
QY	336	CTACTGCAA-----GAGCAATGCTGCAAGAACCAACCAAGATTTCATCCGAAGATC	389		
Db	375	CCACTGCAAACTGTGACATCGCTTGATCCACCAAGTACCAACAGGCAAACTG	434		
QY	390	GTCCACCTTCAGAACTTAGGCAAACTTGTCTATATACCTACCTAGGTACAGTACATCA	449		
Db	435	CAGCAGCTAGTGAAATGATCTCTGATTCATTCATATGATGCTGGGAGCTCTTC	494		
QY	450	AGGAATCTTAGGCTATGATACCGTCACTGTCTCAAC-----	486		
Db	495	CGGTACCTGAGCAGACGACATGTGTGGTCCCTGCAAGTCAAGGTGTGAGCTCTGC	554		
QY	487	-----ATTGTGACATTCAACAGACAGTGAAGCTTGAACCCAAAGACAGGTGA	536		
Db	555	CTGGGCGGTGTCAAAAGTGAAGAGGAGAGGCTTTGGGAGAGCCACCAAGCAGCCAGAT	614		
QY	537	TGTCTTCACTATGAGAAATCGATGGCATCTGTGTATGATACCATGCTGCGGTC	596		
Db	615	CACCTTATGCGAGCAAGTGTGATGAGCATCTGGGAGGCTTACCCCGCATCTCGT	674		
QY	597	AGAGTACTGATTAACCTGTGTTTGAACAATGATGAACCAACCTAGTATGCTCAAGATT	656		
Db	675	CAACAAGTGTGCGCCGCTTTCGACAACTGATGAGAGAAAGCTGTGAGACCAAGAT	734		
QY	657	GTTCTCGTTTAACTGACAGAA-----TGGCCAGAGAGCATGCTCAGCTTGG	707		
Db	735	CTTCTCCTTCTTACCTGAGCAGGAGCCAGATGCGAGCTTGGGAGTGAAGTCTGGG	794		
QY	708	AGCTATTATCCATCTCTACAGAGATCTCTTCACTGCGGTTCCAGTCACTGTGACGA	767		
Db	795	TGGCACAAGCTCAAGATTTACAAGGTTCTGTCTTCACTGAATGACCCGCAAGGC	854		
QY	768	GTACTGCAATTCACCTGAGCAGTGTCCACCATCAGCGGTGTTGTGATGGAAG	827		
Db	855	CTACTGAGAGTCCACCTGAGACAGGTGAGAGTGGCCAGCGGCTGACCTGTGCAAGA	914		
QY	828	TGATGTCAAGCTATCTTGAATACCGTACGTTCAAGCTGTGAGCCTTACAGGACAT	887		
Db	915	GGGCTGTGAGGCAATTGAGACAGGCACTTCCCTCATGATGGGCGGTGATGAGGT	974		
QY	888	TCTCAACTTCAAGAGTATTGAGACCAACCAAGCAAGTATGATGTTGACATAGA	947		
Db	975	GCGGAGCTGCAAGAGGCAATCGGGCCGTGCTGATTCAGGGCGGATCATATATCC	1034		
QY	948	TTGGACAACCTTAGCTACATGCTTACAGTTGTCTTGAATCAACGGAAGATGATCCC	1007		
Db	1035	CTGTGAGAGGTGTTCACCTGCGCGCATCACTGAAGCTGGAGGCAAGGCTTACAA	1094		
QY	1008	ACTGACCCCTCCGCTTATACAGCCAG	1036		
Db	1095	GCTGTCCCAAGAGACTACGCTCAAG	1123		

Search completed: February 5, 2004, 05:20:15
Job time : 102 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2004, 03:08:18 ; Search time 2406 Seconds
(without alignments)
1795.886 Million cell updates/sec

Title: US-09-643-755B-1

Perfect score: 1173
Sequence: 1 atgactctcctaagctctt.....ggctagctaacgactctga 1173

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	894.8	76.3	2733	11	US-09-893-525-6
2	360.8	30.8	1386	12	US-09-971-392-41
3	359.2	30.6	1131	12	US-10-074-978A-7
4	359.2	30.5	1688	12	US-10-074-978A-3
5	357.6	30.5	1131	12	US-10-074-978A-5
6	357.6	30.5	1131	12	US-10-074-978A-11
7	356	30.3	1131	12	US-10-074-978A-9
8	253	21.6	2201	13	US-10-295-027-178
9	248.4	21.2	1366	10	US-09-969-708-441
10	248.4	21.2	1366	10	US-09-969-708-441
11	248.4	21.2	1366	10	US-09-969-708-441
12	248.4	21.2	1366	10	US-09-969-708-441
13	248.4	21.2	1366	10	US-09-969-708-441
14	230.6	19.7	1270	12	US-10-051-874-53
15	226.8	19.3	1173	12	US-10-274-639-30

16	150.6	12.8	1341	13	US-10-220-083-1	Sequence 1, Appli
17	149.2	12.7	1530	10	US-09-887-576-795	Sequence 795, App
18	146.6	12.5	1299	14	US-10-094-080-4	Sequence 4, Appli
19	146.6	12.5	1365	15	US-10-288-222A-29	Sequence 29, Appli
20	146.6	12.5	1425	10	US-09-964-899-10	Sequence 10, Appli
21	146.6	12.5	2061	13	US-10-180-719-16	Sequence 16, Appli
22	144.8	12.3	1521	10	US-09-938-842A-2245	Sequence 2245, Ap
23	144.8	12.3	1521	12	US-09-938-842A-2245	Sequence 13, Appli
24	141.6	12.1	1360	9	US-09-789-919-13	Sequence 43, Appli
25	141.6	12.1	1360	9	US-09-789-919-13	Sequence 2040, Ap
26	138.8	11.8	1234	15	US-10-106-698-2040	Sequence 25, Appli
27	138.2	11.8	2038	13	US-10-101-510-25	Sequence 172, App
28	138.2	11.8	2076	15	US-10-084-817-172	Sequence 5, Appli
29	138	11.8	1828	13	US-10-339-351-5	Sequence 27, Appli
30	136.6	11.6	1988	13	US-10-007-968A-128	Sequence 142, App
31	129.4	11.0	1194	15	US-10-168-425-27	Sequence 27566, A
32	128.2	10.9	600	12	US-10-255-536-142	Sequence 977, App
33	126.6	10.8	1320	15	US-10-369-493-27566	Sequence 4, Appli
34	126.4	10.8	357	13	US-10-125-968-977	Sequence 492, Appli
35	124.6	10.6	1784	13	US-10-339-351-4	Sequence 35234, A
36	123.8	10.6	893	9	US-09-770-445-492	Sequence 46157, A
37	121.4	10.3	1784	13	US-10-339-351-6	Sequence 278, Appli
38	120.4	10.3	480	11	US-09-918-995-35234	Sequence 10, Appli
39	119.4	10.2	1218	12	US-10-369-493-46157	Sequence 43, Appli
40	117.4	10.0	636	9	US-09-925-297-278	Sequence 1, Appli
41	112.2	9.6	2107	15	US-10-206-619-2	Sequence 3, Appli
42	111.6	9.5	1041	13	US-10-324-131-10	
43	108.8	9.3	1420	13	US-09-971-392-43	
44	104.6	8.9	1096	13	US-10-322-746-1	
45	104.2	8.9	1119	13	US-10-322-746-3	

ALIGNMENTS

RESULT 1
US-09-893-525-6
; Sequence 6, Application US/09893525
; Publication No. US20030126631A1
; GENERAL INFORMATION:
; APPLICANT: Moloney, Maurice M.
; TITLE OF INVENTION: Preparation of Heterologous Proteins on Oil Bodies
; FILE REFERENCE: 9369-172
; CURRENT APPLICATION NUMBER: US/09/893,525
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/210,843
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: US 08/846,021
; PRIOR FILING DATE: 1997-04-25
; PRIOR APPLICATION NUMBER: US 08/366,783
; PRIOR FILING DATE: 1994-12-30
; PRIOR APPLICATION NUMBER: US 08/142,418
; PRIOR FILING DATE: 1993-11-16
; PRIOR APPLICATION NUMBER: US 07/659,835
; PRIOR FILING DATE: 1991-02-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 2733
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oleosin-Chymosin Fusion
; NAME/KEY: CDS
; LOCATION: (850)..(1203)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (1444)..(2727)
; OTHER INFORMATION:
; US-09-893-525-6
Query Match 76.3%; Score 894.8; DB 11; Length 2733;

Best Local Similarity 88.4%; Pred. No. 1.2e-302;
Matches 971; Conservative 0; Mismatches 127; Indels 0; Gaps 0;
QY 76 GCTGAGATCAACCCGATTCCTCTCTCAAAAGTAAGTCTCCGTAAAGCGCTGAAGAA 135
Db 1630 GCTGAGATCAACCCGATTCCTCTCTCAAAAGTAAGTCTCCGTAAAGCGCTGAAGAA 1689
QY 136 CATGACTTCTGAAGAAGCTTTGTCAGAAACAAGTAAGGATCAGACGAAGTACTCC 195
Db 1690 CATGAGCTTCTGAAGAAGCTTTGTCAGAAACAAGTAAGGATCAGACGAAGTACTCC 1749
QY 196 GGGTTCGGTGAAGTGTGCTAGCGTCACTTACCACTACCTTATATGTAATCTTTGGG 255
Db 1750 GGGTTCGGTGAAGTGTGCTAGCGTCACTTACCACTACCTTATATGTAATCTTTGGG 1809
QY 256 AAGATCTACCTCGAAACCCCGCTCAAGAGTTCACGGTCTCTTTGATATCTGTTCTCT 315
Db 1810 AAGATCTACCTCGGAGCCCGCCCGAGAGTTCACGGTCTGTTTGAACAATGCTCTCT 1869
QY 316 GACTTCTGGGTTCCCTCTATCTAATGCAAGCAATGCTCGCAAGAACCAAGATTC 375
Db 1870 GACTTCTGGGTTCCCTCTATCTAATGCAAGCAATGCTCGCAAGAACCAAGCGCTTC 1929
QY 376 GATCCGAGAAAGTGTCCACCTTCAGAACTTAGGCAAAACCTTGTCTATACCTACGGT 435
Db 1930 GATCCGAGAAAGTGTCTCACTTCAGAACTTAGGCAAGCCCTGTCTATCCACTACGG 1989
QY 436 ACAGGTAGCATGCAAGGAATCTTAGGCTATGATACCGTCACTGCTCCCAATTGAGAC 495
Db 1990 ACAGGACATGCAAGGCACTCTGGGCTATGACCGTCACTGCTCCCAATTGAGAC 2049
QY 496 ATTCAACACAGCATGAGACTTAGCAACCAAGAACCAAGTATGTTCTTACCTATGCAAA 555
Db 2050 ATTCAACACAGCATGAGACTTAGCAACCAAGAACCAAGTATGTTCTTACCTATGCGAA 2109
QY 556 TTGATGAGGATCTCTGGTATGATACCATCGCTGCGTCAAGATCTGATACCTGTC 615
Db 2110 TTGATGAGGATCTCTGGGATGATACCATCGCTGCGTCAAGATCTGATACCTGTC 2169
QY 616 TTTGACAACATGATGAACCCGACACTAGTACTCAAGACTTGTCTCGTTTACATGAGAC 675
Db 2170 TTTGACAACATGATGAACCCGACACTAGTACTCAAGACTTGTCTCGTTTACATGAGAC 2229
QY 676 AGGAATGCCAGAGAGCATGCTCAAGCTTGAAGCTATTGATCCATCTTACTACACAGGA 735
Db 2230 AGGAATGCCAGAGAGCATGCTCAAGCTTGAAGCTATTGATCCATCTTACTACACAGG 2289
QY 736 TCTCTTCACTGGGTTCCAGTCACTGTGCAAGATCTGGAATTCATCTGAGCAAGTTC 795
Db 2290 TCTCTGACTGGGTGCCCCGAGACAGTCAAGTCTGGAAGTCTGAGCAAGTTC 2349
QY 796 ACCATGACGGTGTGTTGTTGATGTAAGTGAATGTCAGTATCTTGAATACCGGT 855
Db 2350 ACCATGACGGTGTGTTGTTGATGTAAGTGAATGTCAGTATCTTGAATACCGGT 2409
QY 856 AGCTCCAGCTGTGCGACCTAGACGCAATTCCTCAACTTCAGCAAGCTATTGAGCC 915
Db 2410 AGCTCCAGCTGTGCGGCCAGAGGCAATTCCTCAACTTCAGCAAGCTATTGAGCC 2469
QY 916 ACACAGAACAGTACGGTGAATTTGACATGATTGGCAACAACCTTACTACTATGCTTACA 975
Db 2470 ACACAGAACAGTACGGTGAATTTGACATGATTGGCAACAACCTTACTACTATGCTTACA 2529
QY 976 GTTGTCTTTGAGATCAACGCGCAAGATGATCCACTGACCCCTTCGCTTATACAGCCAG 1035
Db 2530 GTTGTCTTTGAGATCAATGCGCAAAATGTACCACTGACCCCTTCGCTTATACAGCCAA 2589
QY 1036 GATCAAGGGTTCTGCAACCAAGTGAATTCAGAGTGAACAATTCGCAAAATGATCTTG 1095
Db 2590 GATCAAGGGTTCTGTACCAAGTGAATTCAGAGTGAACAATTCGCAAAATGATCTTG 2649
QY 1096 GAGAGTGTTCATTTGATGATCTACAGGCTTTTGAACAGGCGCAACCTCTGTTGG 1155

Db 2650 GGGGATGTTTTCATCCGAGAGTATTACAGCGTCTTTGACAGGGCCAAACACTGTGGGG 2709
QY 1156 CTAGCTAAAGCAATCTGA 1173
Db 2710 CTGGCCAAAGCCATCTGA 2727
RESULT 2
US-09-971-392-41/c
; Sequence 41, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Pearson, Cecelia I.
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971,392
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 41
; LENGTH: 1386
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Template ID: 242275.7
US-09-971-392-41
Query Match 30.8%; Score 360.8; DB 13; Length 1386;
Best Local Similarity 60.4%; Pred. No. 3.5e-115;
Matches 662; Conservative 0; Mismatches 407; Indels 27; Gaps 3;
QY 92 TTCTCTTCAACAAGTAAGTCTCTCCGTAAAGCGCTGAAGAAACAATGACTTCTGAAG 151
Db 1288 TCCCTCTTCAACAAGTAAGTCTCTCCGTAAAGCGCACTCTGCGAGCGTGTGCTGAAG 1229
QY 152 ACTTCTTGCAAAACAAGTATGGCATGACAGCAAGTACTCGGCTTCGTGAAGTTG 211
Db 1228 ACTTCTTGAAAGACCAACCTTCAACCGCAAGAAATGTTCTCCCACTGGAAGCTC 1169
QY 212 CTAGCGT-----GCCATTACCAACTTACTGATCTGAATCTTGGGAAGA 259
Db 1168 CCACTCGTGAATGAAGAACAGCCCTGAGAACTACTGATGATGATGATCTTGGCACTA 1109
QY 260 TCTACCTCGGAACCCCGCTCAAGATTCACCGTCTCTTTGATCTGTTCTCTGACT 319
Db 1108 TCGGATTCGGAATCTCTGCGCCAGAGTTTCACTGCTCTTTGACACCGGCTCTCAACC 1049
QY 320 TCTGGGTCCCTCTATCTAATGCAAGCAATGCTGCAAGAACCAAGATTCGATC 379
Db 1048 TGTGGTCCCTCTAGTCTAATGCTTCTCAAGTCTTGTCTGACCAACCAACCGCTTCAAC 989
QY 380 CGAAGAAAGTCTCACTTCCAGACTTAGCAATTAAGCAAACTTGTCTATACCTAGGTAAG 439
Db 988 CTGAGATTCCTTCACTACAGTACAGTTCACAGGAGAGCAAGTCTCATCACTACGACCG 929
QY 440 GTAGCATGCAAGAAATCTTAGGCTATGATACCTGCTACTGTCTCAACATTTGAGACTTC 499
Db 928 GAGCATGACAGCAATCTCGATATACAGACTCTGCTTCAAGTGTGAGGACTCTGACACCA 869
QY 500 AACAGACAGTGAAGCTTAGACCAAGAAACAGAGTATGCTTCACTATGCAAGAAATTCG 559
Db 868 ATCAAGTCTTGGCTTGAAGGAGAGAGCAACCTGCTCTCTCTGTATATGCTCCCTTGG 809
QY 560 ATGCAATCTTGTGATGATACCAATTCGCTCGCTCAAGATCTAGATACCTGTTGTTG 619
Db 808 ATGCAATCTTGGGCTGAGGCTTACCAAGATTTCTCTCTCGGAGGCAACCGCTCTTTG 749
QY 620 ACAACATGATGAACGACACTTAGTACTCAAGACTTGTCTCGTTTACATGGAACAGA 679

Db 748 ACAACATCTGGAACCAAGGCGCTGGTTTCTCAGAACCTCTCTGCTTACCTCAGCGCG 689
 Qy 680 ATGGCGAG--GAGAGCATGCTCAAGCTTGAAGCATATCCATCCATACACAGAT 736
 Db 688 ATGACCAAGTGGCGCGCTGGTGTATCTTTGTGGGATGACTCTTCTTACTACACTGGA 629
 Qy 737 CTCTTCACTGGGTTTCCAGTCACTGTCAGCAGTACTGSCAAATTCAGTGAAGTGC 796
 Db 628 GTCTGAATCTGGGTGCTTCTTACCGTCAGGGGTACTGSCAATCAGCTGAGACATCA 569
 Qy 797 CCATCAGCGGTGTGGTGTTCATGTGAAGTGAATGTCAAGCTATCTTGAATCCGTA 856
 Db 568 CCATGAACGAGAGGCGCATCCGCTGCGCTGAGGGGTGCGAGGCGCATGTGTGACACGGCA 509
 Qy 857 CGTCAAGCTGTGGAGCTGACGAGCGGCAATTTCTAATTCAGACACTATTTGAGCCA 916
 Db 508 CCTCTGCTTACCGGCGCCCAACGAGCCCATTTGCCAATCCAGACGACATCCGAGCCA 449
 Qy 917 CACAGAACCAAGTACGAGTGTTCATAGATATGCGACACCTTAGCTACATGCTACAG 976
 Db 448 GCGAGAACTCAGATGCGACATGTGTGCTGAGTCTGCTCAGCATCAGCAGCTGCGGACA 389
 Qy 977 TTGCTTTGAGATCAACGCGCAAGATGTACCACTGACCCCTCCGCTTATACAGCCGAG 1036
 Db 388 TCGCTTACCATCAATGAGAGTCCAGTACCCCGTCCAGCCAGTGCCTTACATCTCGAGA 329
 Qy 1037 ATCAAGGCTTTCGACCGCTGATTCAGAGTGAAGAACCTTCC-----CAGA 1084
 Db 328 GCGAGGAGGCTGCTCAGTGTGCTTCCAGGGCATACCTCCCAAGCATCTGAGAGC 269
 Qy 1085 AATGATCTTGGGAGATGTGTTTCAATTCGTGATCTAGAGGCTTGTGACAGGGCGACA 1144
 Db 268 TTGGATCTTGGGTATGTCTTCAATCCGCCAGTACTTTCAGCTTTCAGACGGGAAACA 209
 Qy 1145 ACCTGCTGGGCTAGC 1160
 Db 208 ACCAGGTGGGCTGSC 193

RESULT 3
 US-10-074-978A-7/c
 ; Sequence 7, Application US/10074978A
 ; Publication No. US20040010119A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Leite, Mario
 ; APPLICANT: Spytek, Kimberly A
 ; APPLICANT: Guo, Xiaojia (Saeha)
 ; APPLICANT: Fernandes, Elma
 ; APPLICANT: Li, Li
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Liu, Xiahong
 ; APPLICANT: Casman, Steele
 ; APPLICANT: Boldog, Ferenc
 ; APPLICANT: Paturajan, Meera
 ; APPLICANT: Ballinger, Robert
 ; APPLICANT: Vermet, Corine
 ; APPLICANT: Tchernev, Velizar T
 ; APPLICANT: Malyankar, Uriel M
 ; APPLICANT: Gusev, Vladimir
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Mezes, Peter S
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Heyes, Melvin P
 ; APPLICANT: Herman, John
 ; APPLICANT: Pena, Carol E A
 ; APPLICANT: Shinkets, Richard A
 ; APPLICANT: Taupier Jr, Raymond J
 ; APPLICANT: Moore, No. US20040010119A11le
 ; APPLICANT: Shenoy, Suresh
 ; APPLICANT: Edinger, Shlomit
 ; APPLICANT: Gunther, Erik

APPLICANT: Stone, Dave
 APPLICANT: Millet, Isabelle
 APPLICANT: Peyman, John
 APPLICANT: Smithson, Glenda
 TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 FILE REFERENCE: 21402-269
 CURRENT FILING DATE: 2003-01-07
 PRIOR APPLICATION NUMBER: US/10/074, 978A
 PRIOR FILING DATE: 2001-02-12
 PRIOR APPLICATION NUMBER: 60/335,109
 PRIOR FILING DATE: 2001-10-31
 PRIOR APPLICATION NUMBER: 60/312,284
 PRIOR FILING DATE: 2001-08-14
 PRIOR APPLICATION NUMBER: 60/268,496
 PRIOR FILING DATE: 2001-02-13
 PRIOR APPLICATION NUMBER: 60/276,703
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/330,293
 PRIOR FILING DATE: 2001-10-18
 PRIOR APPLICATION NUMBER: 60/322,127
 PRIOR FILING DATE: 2001-11-21
 PRIOR APPLICATION NUMBER: 60/280,899
 PRIOR FILING DATE: 2001-04-02
 PRIOR APPLICATION NUMBER: 60/310,797
 PRIOR FILING DATE: 2001-08-08
 PRIOR APPLICATION NUMBER: 60/268,646
 PRIOR FILING DATE: 2001-02-14
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 547
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 7
 LENGTH: 1131
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-074-978A-7

Query Match 30.6%; Score 359.2; DB 12; Length 1131;
 Best Local Similarity 60.3%; Pred. No. 1.1e-114;
 Matches 661; Conservative 0; Mismatches 408; Indels 27; Gaps 3;

Qy 92 TTCCCTCTCAAAAGTAAAGTCTCCGTAAGCGCTGAAGAAACATGACTTCTAGAG 151
 Db 1112 TCCCTCTATAGAAAGAGTCTTGAAGGCGCACCTGTCCGAGGTGGCTGTGAAG 1053
 Qy 152 ACTTCTTGCAAAACAGATGAGCATGAGAGCAAGTACTCCGGCTTGGGTAATTG 211
 Db 1052 ACTTCTTGAAAGAACACAACTCAACCCAGCAGAAAGTACTTCCCGAGTGGAGGCTC 993
 Qy 212 CTAGCGT-----GGCACTTACCACTACCTTGATATGATCAATCTTGGAGAGA 259
 Db 992 CCACCTGTGATGATGAACAGCCCTGAGAACTACCTGATATGAGATCTTGGCACTA 933
 Qy 260 TCTACTGGAACCCCGCTCAAGAGTTCAACGTTCTTTGATACGTGGTCTCTGACT 319
 Db 932 TCGCATGGAACCTCTGCGCAGAGATTTCAACCGTCTTTGACACGGCTCTCTCAAC 873
 Qy 320 TCTGGGTTCCCTATCTACTGCAAGAGCAATGCTGCAAGAACCAACCAAGATTGATC 379
 Db 872 TGTGGGTCCCTCAGTCTACTGCTCAGGTCTTGCCTGCAACCAACCAACCGCTTCAAC 813
 Qy 380 CGAAGAACTGCTCAACCTTCAAGAACTTGAAGCAACCTTGTCTATACATACGTATAG 439
 Db 812 CTGAGATCTTTCACCTTCAACGATCCACAGCGAGACAGTCTTCAATCACTTACGGACCG 753
 Qy 440 GTAGATGCAAGAACTTGAAGTATGATACCTGATCTTCAACATTTGAGCATTC 499
 Db 752 GAGGATGAGAGGATCTCGGATAGCACTGTCAAGTTGAGGAGATCTCTAGACCA 693
 Qy 500 AACGACATGAGACTTGAACCAACCAAGACAGATGATCTTCACTATGACAGATTG 559
 Db 692 ATCAGATCTTGGGCTGAGGAGAGAGCACTGGCTCTTCTCTATATGCTCTCCCTTGG 633

```
QY 560 ATGGCATCTTGTGATGCAATACCATGCTGCGTCAAGATCTGATACCTGTGTTG 619
Db 632 ATGGCAATCTGGGGCTGGCTACCCAGCATTTCTCTCCGGGCGACACCCGTTTGG 573
QY 620 ACAACATGATGAAACGACACCTAGTACCTCAAGCTTGTCTGTGTTTACATGAGAGA 679
Db 572 ACAACATGATGAAACGAGGCTGTGTTTCTCAGAACCTCTTCTGTCTACCTCAGGCCG 513
QY 680 ATGGCCAG--GAGAGATGCTCAACGCTTGAGCTTATGATTCATCTTACATCAAGAT 736
Db 512 ATGACCAAGATGCGAGGTGTGATCTTTGTGTGCAATGACTCTTCTTACATCTGAA 453
QY 737 CTCTTACAGGGGTTCAGTACCTGTGAGAGATGACAGGCAATTCAGCTGTGAGAGTCA 796
Db 452 GTCTGAATCTGGGTGCTGTTCACCTGAGGGTTATCTGGCAGATCACCGTGAACATCA 393
QY 797 CCATGACGGGTGTGTTGTTCATGATGAAAGTGGATGATCAAGCTATCTTGATACCGTGA 856
Db 392 CCATGAACGAGAGGCGCATGCTGCGCTGAGGGCTGCCAGGCTTGTGACACCGGCA 333
QY 857 CGTCAAGCTGTGTGCGACCTAGACGACCATTTCTCAACATTCAGCAAGCTATTGGAGCCA 916
Db 332 CCTCTGTGTCGACCGGCGCCCAACGAGCCCATGTGCCAATCCAGAGCGACATGCGAGCCA 273
QY 917 CACAGAACCAAGTACGGTGAATTTGACATGATTTGCGACCAACTTAGTACATGCTTACAG 976
Db 272 GCGAGAACTCAGATGGGAGATGTGTGTGACGTGCTCAGCATGAGAGCTGCTCCGACA 213
QY 977 TTGTCTTTGAGATCAACGCAAGATGATGACCTGACCCCTCGCTTATACAGGACGAG 1036
Db 212 TGTCTTTCACCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 153
QY 1037 ATCAAGGGTTCTGACACCATGATGATTCAGAGTGAAGCAATTC--CAGA 1084
Db 152 GCGAGGGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 93
QY 1085 AATGATCTTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1144
Db 92 TTTGGATCTCTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 33
QY 1145 ACCTCGTTGGGCTAGC 1160
Db 32 ACCAGGTGAGCTGTGC 17

RESULT 4
US-10-074-978A-3
; Sequence 3, Application US/10074978A
; Publication No. US20040010119A1
; GENERAL INFORMATION:
; APPLICANT: Leite, Mario
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Fernandes, Elma
; APPLICANT: Li, Li
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Liu, Xiaohong
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Patlurajan, Meera
; APPLICANT: Bialock, Angela
; APPLICANT: Ballinger, Robert
; APPLICANT: Vermet, Corine
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Gusev, Vladimir
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter S
; APPLICANT: Elleman, Karen
; APPLICANT: Heyes, Melvin P
; APPLICANT: Herrman, John
; APPLICANT: Pena, Carol E A
; APPLICANT: Shimkets, Richard A
```

```
APPLICANT: Taupier Jr, Raymond J
APPLICANT: Moore, No. US20040010119A111e
APPLICANT: Shenoy, Suresh
APPLICANT: Edinger, Shlomilt
APPLICANT: Gunther, Erik
APPLICANT: Stone, Dave
APPLICANT: Millet, Isabelle
APPLICANT: Peyman, John
APPLICANT: Smithson, Glenda
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-269
CURRENT APPLICATION NUMBER: US/10/074, 978A
CURRENT FILING DATE: 2003-01-07
PRIOR APPLICATION NUMBER: 60/268, 221
PRIOR FILING DATE: 2001-02-12
PRIOR APPLICATION NUMBER: 60/335, 109
PRIOR FILING DATE: 2001-10-31
PRIOR APPLICATION NUMBER: 60/312, 284
PRIOR FILING DATE: 2001-08-14
PRIOR APPLICATION NUMBER: 60/268, 496
PRIOR FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/276, 703
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/330, 293
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/322, 127
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 60/280, 899
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 60/310, 797
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: 60/268, 646
PRIOR FILING DATE: 2001-02-14
Remainder Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 547
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 1688
TYPE: DNA
ORGANISM: Homo sapiens
US-10-074-978A-3

Query Match 30.6%; Score 359.2; DB 12; Length 1688;
Best Local Similarity 60.3%; Pred. No. 1.4e-114;
Matches 661; Conservative 0; Mismatches 408; Indels 27; Gaps 3;

QY 92 TTCTCTTCAAAAGTAAAGTCTCTCGTAAAGGCGCTGAAGAAATGAGACTTCTGAAG 151
Db 412 TCCCTCTCAATCAAGAAAGTCTCTTGAAGGCGACCTGTCCAGAGCTGCTGTAAG 471
QY 152 ACTTCTTGCAAAACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 211
Db 472 ACTTCTTGAAAGACCAACCTCAACCCAGCAAGAAAGTATCTTCCAGTGGAGGCTC 531
QY 212 CTAGCGT-----GCCACTTACCACTACTCTTATATGCTCAATCTTTGGAGAAG 259
Db 532 CCACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 591
QY 260 TCTAACCTGGAAACCGGCTCAAGAGTACCGTCTCTTTGATATGTTGCTCTGAGCT 319
Db 592 TCGGCAATCGGAATCTCTGCGCAGATTTCACTGTCTCTTTGACACCGGCTCTCAAC 651
QY 320 TCTGGTTCCTCTATCTACTGCAAGAGATGCTCTGCAAGAACCAACCAAGATTGAGATC 379
Db 652 TGTGGGTGCTCTAGTCTACTGTCTCAAGTCTTCTGCAACCAACCAACCGCTTCAAC 711
QY 380 CGAGAAAGTGTCTCACTTTCAGAACTTGAAGCAACCTTGTCTATACATGATGATG 439
Db 712 CTGAGATTTCTTCACTTCAAGGCGACCAAGCAAGCAAGCTTCTCACTTCAAGGACCG 771
QY 440 GTAGCATGCAAGAACTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATG 499
Db 772 GAGCATGACAGGATCTCTGGATGACGACCTGTCAAGTTGAGGCAATCTTGAACCA 831
```



```

Db      812 CTGAGAGTTCTTCCACTACCACTCCACGAGAGACAGTCTTCATCCTACCTACCGGACCG 753
Qy      440 GTAGCATGCAAGAAATCTTAGGCTATGATACCGCTACCTGTCTCCAACTTTGTGACATTC 499
Db      752 GCAGCATGACAGGATCTCTCGATATGACACTGTCCAGGTTGAGGAGATCTTGACACCA 693
Qy      500 AACAGACAGTAAAGATTAAGACCCAGAACAGATGATGTCTTCACTATGCAAGATTG 559
Db      692 ATGAGATCTTCCGCTGAGGAGAGACGAACTGCTCTTCTCTGATTAATGCTCCCTTCG 633
Qy      560 ATGGCATCTCTGATATGATACCACTCCGCTGAGAGATCTGATACCTGATTTG 619
Db      632 ATGGCATCTCGGGGCTGCTTACCCAGCATTTCTCTCCGGGGCCACACCGCTCTTG 573
Qy      620 ACAACATGATGAACCGACACCTAGATCTCAAGCTTTGTTCTGCTTTTACATGACAGGA 679
Db      572 ACAACATCTGGAACCGAGGCGCTGTTCTCAGAGACCTTCTCTGTCTACCTCAGGCGG 513
Qy      680 ATGGCCAG---GAGAGATGCTCAGCGTTGAGCTATGATCCATCTACTACACAGGAT 736
Db      512 ATGACCAAGATGAGGAGGATGATCTTGTGTGATGATCTTCTTACTTACACTGGA 453
Qy      737 CTCTTCACTGGGTTCCAGTCACTGTGACAGATCTGGCAATTCATCTGAGACATGTC 796
Db      452 GTCTGAATCTGGTGCTGCTTACCGTCAAGGTTACTGGACATCACCGTGAACATCA 393
Qy      797 CCATCAGCGGTGTGTGTTGATGATGAAAGTGTGATGATCTATCTTGATACCGGTA 856
Db      392 CCATGAACGAGAGGCGATCGCTCGCTGAGGAGGCTCCAGGCACTGTTGACACGGGCA 333
Qy      857 CGTCCAGCTGTGTGCGACCTTACAGAGGCAATTTCTCAACATTCAGCAAGCTATTGAGCCA 916
Db      332 CCTCTGCTGACCGGCGCAACAGCCCATGTGCAACATTCAGAGGAGCATCGAGCCA 273
Qy      917 CACAGAACAGTACGAGTGAATTTGACATGATTTGCGAACACTTAGCTATCATGCTTACAG 976
Db      272 GCGAGAACTGATGAGGAGATGTGTCTCACTCTCAGCATCAGACATGAGGCTGCGGACA 213
Qy      977 TTGTCTTGAATCAACGAGAAATGTACCACTGACCCCTCCGCTTATACAGCCAGG 1036
Db      212 TCGCTTCAACATCAATGAGATCGATCCAGTACCCGCGGACCAAGTCCCTATCCTGAGA 153
Qy      1037 ATCAAGGTTCTGACCAAGTGAATTTCCAGAGTGAAGCAATTC-----CAGA 1084
Db      152 GCGAGGAGGAGCTGATGATGAGTCTTCAGGCGATGAACCTCCCAAGAACTGAGAGC 93
Qy      1085 AATGATCTTGGAGATGTGTTCATTCGTGATCTACAGGCTTTTGAAGGGGCAACA 1144
Db      92 TTGATCTCTGGGATGTCTTCTTCATCCGCAATCTTACCGTCTTGAAGGGCAACA 33
Qy      1145 ACCCTGTTGGGCTAGC 1160
Db      32 ACCAGGTGAGGCTGCG 17

```

RESULT 6
US-10-074-978A-11/c
Sequence 11, Application US/10074978A
Publication No. US20040010119A1

GENERAL INFORMATION:

APPLICANT: Velite, Mario
APPLICANT: Spytek, Kimberly A
APPLICANT: Guo, Xiaojia (Sasha)
APPLICANT: Fernandez, Elma
APPLICANT: Li, Li
APPLICANT: Ketuda, Ramesh
APPLICANT: Liu, Xiaohong
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
APPLICANT: Patuturajan, Meera
APPLICANT: Blalock, Angela
APPLICANT: Ballinger, Robert
APPLICANT: Vermet, Corine

```

APPLICANT: Tchernev, Velizar T
APPLICANT: Malyankar, Uriel M
APPLICANT: Gusev, Vladimir
APPLICANT: Rastelli, Luca
APPLICANT: Mezes, Peter S
APPLICANT: Ellerman, Karen
APPLICANT: Heyes, Melvin P
APPLICANT: Herrman, John
APPLICANT: Pena, Carol E A
APPLICANT: Shimketa, Richard A
APPLICANT: Taupier Jr, Raymond J
APPLICANT: Moore, No. US20040010119A11le
APPLICANT: Shenoy, Suresh
APPLICANT: Edinger, Shlomit
APPLICANT: Gunther, Erik
APPLICANT: Stone, Dave
APPLICANT: Millet, Isabelle
APPLICANT: Peyman, John
APPLICANT: Smithson, Glenda
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-269
CURRENT APPLICATION NUMBER: US/10/074,978A
CURRENT FILING DATE: 2003-01-07
PRIOR APPLICATION NUMBER: 60/268,221
PRIOR FILING DATE: 2001-02-12
PRIOR APPLICATION NUMBER: 60/335,109
PRIOR FILING DATE: 2001-10-31
PRIOR APPLICATION NUMBER: 60/312,284
PRIOR FILING DATE: 2001-08-14
PRIOR APPLICATION NUMBER: 60/268,496
PRIOR FILING DATE: 2001-02-13
PRIOR APPLICATION NUMBER: 60/276,703
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/330,293
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/322,127
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: 60/280,899
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 60/310,797
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: 60/268,646
PRIOR FILING DATE: 2001-02-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 547
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 1131
TYPE: DNA
ORGANISM: Homo sapiens
US-10-074-978A-11

```

Query Match 30.5%; Score 357.6; DB 12; Length 1131;
Best Local Similarity 60.2%; Pred. No. 4.1e-114;
Matches 660; Conservative 0; Mismatches 409; Indels 27; Gaps 3;

```

Qy      92 TTCTCTTCAAAAGGTAAGTCTCTCCGTAAGCGCTGAAGAAACATGCACTTCTAGAAG 151
Db      1112 TCCCTCTATCAAGAAAGATCTCTTGAAGGAGACCTGTCCAGAGCTGCTGCTGAAG 1053
Qy      152 ACTTCTTGCAAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 211
Db      1052 ACTTCTTGAAAGCAACCTTCACCCAGCCAGAAAGTACTTCCCGAGTGGAGGCTC 993
Qy      212 CTAGCGT-----GCCACTTACCACTACTGATGATGATGATGATGATGATGATGATG 259
Db      992 CCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 933
Qy      260 TCTACCTGGAACCCGCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 319
Db      932 TCGGATCGGAATCTCTGCGCAGAGATTTCACTGCTCTTTGACACGGGCTCTCAACC 873
Qy      320 TCTGGGTTCTCTATCTACTGCAAGAGCAATGCTGCAAGAACCAACAAAGATTGATC 379

```


QY	260	TCATCCCTCGAAGACCCCGCTCAAGAGTTCTACCGTCTCTTTGATATCTGGTCTCTAGCT	319
Db	932	TCCGCAATCGAAGACTCTGCCCGAGATTTTCAACCTGCTCTTTGACACCGGCTCTCTCAAC	873
QY	320	TTGGGATTCCTCTATCTACTGCAAGACCAATGCTGCAAGAACCCAAAGATTGCATG	379
Db	872	TGTGGGTGCCCTCAAGTCTACTGCTTCCAGTCTTGGCTTGCTGCAACCAACCAACCGCTTCAAC	813
QY	380	CGAAGAAAGTGTCCACCTTCCAGAACTTAAGCCAAACCTTGTCTATACATACTACGGTACAG	439
Db	812	CTGAGGATCTTCCACTTACCAAGTCACACAGCGAAGACAGTCTCATCACTTAACGGCACCG	753
QY	440	GTACACTGCAAGAACTTTAGAGTATGATACCGGTACCGGTCTCCAACTTGGACATTTC	499
Db	752	GCAGCATGACAGGCATCTCGGATGACGACATCTGCCAGGTGGAGGACATCTTGACACCA	693
QY	500	AACAGACAGTAGAATTAGCAACCAAGAACCAAGATGATGTCTTCACTATGCAAAATTGC	559
Db	692	ATCAGATCTTGGCGCTGAGCGAAGACCGAATCGGTCTCCTGATATTGCTCTCCCTTG	633
QY	560	ATGGCATCTTGTGATATGGCATATCCCATGCTGCGGTCAAGATATCTGATCTGTGTTTG	619
Db	632	ATGGCATCTTGGGGGTGGCTTACCCCAAGCAATTTCTCTCCGGGGACACACCCGCTTTTG	573
QY	620	ACAACATGATGATACCGACACCTTAGTAGCTCAAGACTTGTCTCGGTTTATACATGACAGA	679
Db	572	ACAACATGATGAAACCAAGGCGCTGGTTTCTCAGAACCTCTTCTGTCTTACCTCAGCGCG	513
QY	680	ATGGCCAG--GAGAGCATGTCTACGCTTGGAGCTATTGATTCATCTACTACACAGAT	736
Db	512	ATGACAAAGATGGGACGCTGGTGATCTTTGGGGATTTGACTCTTACTATACATGGAA	453
QY	737	CTCTTCACTGGGTTCCAGTCACTGTGACAGCAATCTGCGCAATTCACTGTGACAGTGTCA	796
Db	452	GTCGAACTGGGTGTCTGTACCTGTGACGGGTTTACTGGCAGATCAACGTTGACACATCA	393
QY	797	CCATCAGCGGTGTGGTTGTGTGATGTAAGGTGATGTCAAGCTATCTTGGATACGGTGA	856
Db	392	CCATGAAACGAAAGACCAATCGCTGTGTGTGAGGGCTGCGAGGCAATGTTGACACCGGA	333
QY	857	CGTCCAAAGCTGTGCGACCTTAGCAGCGCAATTTCTCAATTCCAGCAAGCTATTGGAGCCA	916
Db	332	CCTCTGTGCTAACCGGCGCCAACCGACCCCATTTGCCAATTCAGAGCGACATCGGAGCCA	273
QY	917	CACAGAACAGTACGGTAGTTGACATATGATTTGCGACAACTTAAGCTACATGCTTACAG	976
Db	272	GCGGAAATCCAGATGGCAGACATGTGTGTACGCTGTCAAGCCATCAGACGCTGCCGACA	213
QY	977	TTGTCTTTGAGATTCAAAGGCAAGATGTATCCCATGACCCCGCTCGGCTATATCCAGCCAG	1033
Db	212	TGCTTTTACCAATCATATGAAAGTCCAGTACCCCGTGTCAACCCAGTGTCTTACATCTCGAGA	153
QY	1037	ATCAAGGGTTCTGCAACGATGGATTCCAGAGTGAGAACCATTTCC-----CAGA	1084
Db	152	GCGAGGGGAGGTGATCATGTGGCTTCCAGGGGATTAAGTCCCAACCGAATGTGAAGGC	93
QY	1085	AATGATCTTTGGAGATGTGTTCAATTGTAAGTACTACAGGCTCTTTGACAGGGCCACA	1144
Db	92	TTTGATCTGTGGTATGTCTTCAATCCGCCAGTACTTTAACGCTTTCGACAGGGCAACA	33
QY	1145	ACCTGTTGGGCTAGC	1160
Db	32	ACCAAGTACGCTGGC	17

RESULT 8
US-10-295-027-178
; Sequence 178, Application US/10295052Z
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afari, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.

```

1  APPLICANT:  Glieh, Kurt C.
2  APPLICANT:  Glyme, Richard
3  APPLICANT:  Hevezi, Peter A.
4  APPLICANT:  Mack, David H.
5  APPLICANT:  Murray, Richard
6  APPLICANT:  Watson, Susan R.
7  APPLICANT:  Eos Biotechnology, Inc.
8  TITLE OF INVENTION:  Methods of Diagnosis of Cancer, Compositions and
9  FILE OF INVENTION:  Methods of Screening for Modulators of Cancer
10 FILE REFERENCE:  018501-012500US
11 CURRENT APPLICATION NUMBER:  US/10/295,027
12 CURRENT FILING DATE:  2002-11-13
13 PRIOR APPLICATION NUMBER:  US 09/663,733
14 PRIOR FILING DATE:  2000-09-15
15 PRIOR APPLICATION NUMBER:  US 60/350,666
16 PRIOR FILING DATE:  2001-11-13
17 PRIOR APPLICATION NUMBER:  US 60/335,394
18 PRIOR FILING DATE:  2001-11-15
19 PRIOR APPLICATION NUMBER:  US 60/332,464
20 PRIOR FILING DATE:  2001-11-21
21 PRIOR APPLICATION NUMBER:  US 60/334,393
22 PRIOR FILING DATE:  2001-11-29
23 PRIOR APPLICATION NUMBER:  US 60/340,376
24 PRIOR FILING DATE:  2001-12-14
25 PRIOR APPLICATION NUMBER:  US 60/347,211
26 PRIOR FILING DATE:  2002-01-08
27 PRIOR APPLICATION NUMBER:  US 60/347,349
28 PRIOR FILING DATE:  2002-01-10
29 PRIOR APPLICATION NUMBER:  US 60/355,250
30 PRIOR FILING DATE:  2002-02-08
31 PRIOR APPLICATION NUMBER:  US 60/356,714
32 PRIOR FILING DATE:  2002-02-13
33 Remaining Seq ID Application data removed - See File Wrapper or PALM.
34 NUMBER OF SEQ ID NOS: 1386
35 SOFTWARE:  PatentIn Ver. 2.1
36 SEQ ID NO 178
37 LENGTH: 2158
38 TYPE: DNA
39 ORGANISM: Homo sapiens
40 US-10-295-027-178

```

Query Match	21.6%;	Score 253;	DB 12;	Length 2158;
Best Local Similarity	55.3%;	Pred. No. 3.5e-77;		
Matches	621;	Conservative	0;	Mismatches 460; Indels 42; Gaps 5
QY	87	CCGCATTCCTCTCTACAAAGTAGTCTTCCTCGTAAGCGCTGAAGCAATGCACTTCT	146	
Db	115	CAGGGTGGCCCTCGAGGAGGCATCCGCTCTCAMAAGAAAGCTGGGGGACGAGCAACT	174	
QY	147	AGAAAGCTCTTGAGAAACAACAGTATGGCATGA-----GCAAGCAATGATCCCGGCTT	200	
Db	175	CTCTGAGTTCTTGAGAAATCCCAATAATTGGACATGATCCAGTTCACCGAGTCTGCTCAAT	234	
QY	201	CGGTGAAGTTGCTAGCGTGCACCTTACCACTACCTTGATAGTCATATCTTTGGGAAGAT	260	
Db	235	GGACCAAGTGCACAAGAAACCCCTCATCACTACTGGATATGGAATACTTCGGCATAT	294	
QY	261	CTACCTCGGAACCCCGCTCAAGAGTTCAACCGTCTCTTTATATCTGTTTCTCTGACATT	320	
Db	295	CTCATATGGCCCCCAACCAAGAACTTCACGTATCTTTCACACTGGCTCTCTCAACCT	354	
QY	321	CTGGGTTCCCTCTATCTACTGACAAAGAGCAATGCCGCAAGAACCAACCAAGATTGCATCC	380	
Db	355	CTGGGTCCCTCTCTGTGATCTGACACTAGGCCACGCTGGCAAGCGACACAGAGTTCCAGCC	414	
QY	381	GAGAAAGTCGTCAACCTTCCAGAACTTAAGGCAAAACCTTGTCTATATCACTACGGTACAGG	440	
Db	415	TTCCACAGTCACAGCATPACAGCCAGCGAGTCAATCTTCTTCATCTAGTTMGAAACCGG	474	
QY	441	TAGCATGAAAGAACTTTAGGCTATGTAACCGTCACTGTCTCCAAATTTTGGACATTTCA	500	
Db	475	GAGCTTTGTCGGGATCATTTGAGCGCAAGTCTCTGTGGAAGAGCATCAACCGGTGGG	534	

501 ACAGACAGTAGAGCTTAGACCCCAAGAACAGAGTGTCTTACCTATGAGAAATTGCA 560
 535 CCAGCAGCTTTGGAGAAAGTGTACAGAGCCAGCCGACCTTTGTGATGACAGATTGGA 594
 561 TGGCATCTTGGTATGGCATACCCATCGTCCGCTCAGAGTACTGTGATCTGTGTTTGA 620
 595 TGGAAATCTGGGCTGGGATACCCCTCTTGGCTGTGGAGAGTGAATCCAGTATTGTA 654
 621 CAACATGATGAACCGACACCTAGTAGCTCAAGACTTCTCGGTTTATACAGACGAA 680
 655 CAACATGATGCTCAGAACCTGTGATGCTGCGATGTTTCTGTCTATACAGAGTGA 714
 681 -----TGGCCAGAGAGAGATGCTCAAGCTTGGAGCTATTGATCCATCTACTAC 731
 715 CCAGAGAGTGTGGGGAGGAGAGAGTGAATTTTGGAGGCTAGAACCACTCCATTTTC 774
 732 AGGATCTCTTCACTGGGTTCCAGTCACTGTGCAAGATGCTGCAATTCACTGTGAC 791
 775 TGGGAGCTGAAATGGGTCCTCAGTCAACCAAGCAAGCTTACTGGCAGATTGCACTG 834
 792 TGTCAACATCAGCGGTGTGTGTGATGAGTGAAGTGTGATGATCAATCTTGGATAC 851
 835 CATCAGAGTGGAGAGCACTGTATGTTCTGCTCCAGGGCTGCGAGCCATTGTGACAC 894
 852 CGGTACGTCCAGCTGTGTGCACTAGCAGCAGCAATTTCAATTCAGCAGCTATTGG 911
 895 AGGACTTCCCTCATCTACCTGCGCTTCCAGCAAGATTAAGCAGCTGCAAAAGCCATT 954
 912 AGCCACACAGAAACAGTACGCTGAGTTTGAATGATGCGACAACTTGAAGTACAGCC 971
 955 GCGAGCC---CGGTGATGAGAAATATGCTGTGAGTGTGCAACTTAAAGTACAGCC 1011
 972 TAGAGTGTCTTGAATGAAGTGAAGTGAATCCACAGACCCCTCGGCTATACAG 1031
 1012 GATGTACCTTCACTTAAAGAGTCCCTTATACCTTACGCAAGCTTACAGCCCT 1071
 1032 CCAGATCAAG-----GGTTCTGACACAGTGAATTCAGAGTGAAGAACCATTC 1079
 1072 ACTGACTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131
 1080 CCAGAAA-----TGAATCTTGGAGATGATGATGATGATGATGATGATGATGAT 1127
 1132 CCCTCAGCTGGGCCCCCTGTGATCTGTGGGATGCTTCAATTCAGAGTTTACTCAG 1191
 1128 CTTTGAACAGGCGCAACACTGTGTGGCTAGCTTAAGCAATC 1170
 1192 CTTTGAACGCTGGAAATTAACCGTGTGGAGCTGGCCCGCAGCAGTC 1234

RESULT 9
 US-10-101-510-457
 ; Sequence 457, Application US/10101510
 ; Publication No. US20030148295A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WAN, JACKSON
 ; APPLICANT: WAN, YIXIN
 ; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
 ; FILE REFERENCE: 15117.0012
 ; CURRENT APPLICATION NUMBER: US/10/101,510
 ; PRIOR FILING DATE: 2002-03-20
 ; PRIOR APPLICATION NUMBER: 60/276,947
 ; NUMBER OF SEQ ID NOS: 805
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 457
 ; LENGTH: 2201
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (1432)..(1467)
 ; OTHER INFORMATION: a, t, c, g, other or unknown
 US-10-101-510-457

Query Match 21.6%; Score 253; DB 13; Length 2201;
 Best Local Similarity 55.3%; Pred. No. 3.6e-77;
 Matches 621; Conservative 0; Mismatches 460; Indels 42; Gaps 5;
 87 CCGATTCCTCTTACAAAGGTAAGTCTCTCCGTAAGGCGCTGAAGAAATGACCTTCT 146
 158 CAGGGTGCCCTCAGAGAGCATCCGCTCCCAAGAAAGAGTGGCGGACCGAGCCACT 217
 147 AGAAGCTTCTTGCAGAAACAAAGTATGAGATCA-----GCGAGATCTCCGGCTT 200
 218 CTGTAGTCTTGGAAATCCCATTAATTTGACATGATTCAGATTACCGAGTCTCTCAAT 277
 201 CGGTAAAGTGTGATGCGTACGCACTTACCACTTATGATGATGATGATGATGATGAT 260
 278 GAGCAGAGTGCAGAGAAACCCCTCATACACTTGTGATATGAAATATCTTGGACATAT 337
 261 CTACCTCGAAACCCCGCTCAAGAGTTCACCGTTCTCTTGAATACCTGTTCTCTGACTT 320
 338 CTGATGCTCCCTCCAGCAGAACTTCACTGTATCTTGCACATCTGCTCCCAACT 397
 321 CTGGGTTCCCTCTATCTCTGAGAGCAATGCTGCAAGAACCAACCAAGATTGATTC 380
 398 CTGGGTTCCCTCTGTGTCTGATCTGATGCTGATGCTGATGCTGATGCTGATGCT 457
 381 GAGAAAGTGTGCACTTCCAGAACTTGAAGAAACCTTGTCTATACCTACGTTACAG 440
 458 TTCCAGTTCAGACATACAGCAGCAGCAGTCAATTTCTCATATTCAGATGAAACGG 517
 441 TAGCATGCAAGAAATCTTGAAGCTATGATGATGATGATGATGATGATGATGATGAT 500
 518 GAGCTTGTCCGGAATCATTTGAGAGCGGACCAAGTCTCTGTGAAGAGCACTAACCGT 577
 501 ACAGACAGTAGAGCTTAGACCCCAAGAACAGAGTATGTTCTACCTTATGAGAAATTG 560
 578 CCAGAGTGTGAGAAAGTGTACAGAGCCAGGCGCAGCTTGTGATGAGAGTGAATTTGA 637
 561 TGGCATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 620
 638 TGGAAATCTTGGGCTGGGATACCCCTCTTGGCTGTGGAGAGTGAATCTCAAGTATTGA 697
 621 CAACATGATGAACCGACACTAGTACTCAAGACTTGTCTCGTGTTCATGAGACAGAA 680
 698 CAACATGATGCTCAGAACTGTGTGATCTTGCATGATGATGATGATGATGATGATGAT 757
 681 -----TGGCCAGAGAGAGATGCTCAGCGCTTGAAGCTATGATCATTCTACTAC 731
 758 CCAGAGAGTGTGTGGGAGGAGAGCTGATTTTGGAGGCTAACAGCACTCCATTCTTC 817
 732 AGGATCTTCTACCTGCTTCCAGATCACTGTGCAAGATGATGATGATGATGATGATGAT 791
 818 TGGAGCTTGAATTTGGGTCCTCAGATCAACCAAGCTTATGAGAGTGTGATGATGAT 877
 792 TGTCAACATCAGCGGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 851
 878 CATCAGAGTGGAGAGCACTGTATGTTCTGCTCCAGAGGCTGCGAGCCATTGTGACAC 937
 852 CGGTACGTCCAGCTGTGTGCACTAGCAGCAGCAATTTCAATTCAGCAGCTATTGG 911
 938 AGGACTTCCCTCATCTACCTGCGCTTCCAGAAAGTGAAGAGCTGCAAAAGCCATTGG 997
 912 AGCCACACAGAAACAGTACGCTGATGATGATGATGATGATGATGATGATGATGATGAT 971
 998 GCGAGCC---CGGTGATGAGAAATATGCTGTGAGTGTGCAACCTTAAAGTATGCC 1054
 972 TACAGTGTCTTGTGATGATCAAGGCAAGATGATCACTGACCCCTCGCTATACAG 1031
 1055 GATGTACCTTCACTTCACTTAAAGAGTCCCTTATACCTTACGCCCAATGCTTACACCT 1114
 1032 CCAGATCAAG-----GGTTCTGACACAGTGAATTCAGAGTGAAGAACCATTC 1079
 1115 ACTGACTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1174

```
QY 1080 CCAGAAA-----TGGATCTTGGAGAGATGCTTGCATTCGTGAGTACTACAGCCT 1127
Db 1175 CCTCCAGCTGGGCCCCCTCTGAGATCTCGGAGAGATCTTCAATTCAGAGTTTACATCAGT 1234
QY 1128 CTTTGAAGGGCCACCAACCTCGTGGGCTTAGCTAAAGCAATC 1170
Db 1235 CTTTGAACCGTGGGAATTAACGCTGTGGGACTGGCCCCAGCAGTC 1277

RESULT 10
US-09-969-708-441
; Sequence 441, Application US/09969708
; Patent No. US20020102532A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-70
; CURRENT APPLICATION NUMBER: US/09/969,708
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: US/60/237,606
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,608
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,425
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 441
; LENGTH: 1366
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-708-441

Query Match 21.2%; Score 248.4; DB 10; Length 1366;
Best Local Similarity 53.8%; Pred. No. 1.1e-75;
Matches 594; Conservative 0; Mismatches 486; Indels 24; Gaps 3;

QY 92 TTCTCTCTCAAAAGTAAGTCTCTCCGTAAAGCGCTGAAGAAACATGACCTTCTAGAAG 151
Db 1112 TGGCCCCGAAAGAAATTTAAGTCTATCCGTAGACCATGAAGAGAAAGGCTTGTGGGGG 171
QY 152 ACTTCTTGAGAAACACAGATGATGCAATGAGCAAGTATCCTGGGCTTGGTGAAGTTG 211
Db 172 AGTTCCTGAGAGACCCACAGATGATCTCTGCTTGAAGTACCGCTTTGTGACCTCAGCG 231
QY 212 CTAGCGTGCACCTTACCACTACCTTGTATGATCAATCTTTGGAGAATCTACCTCGAA 271
Db 232 TGACCTACGAGCCCATGCGCTACATGATGCTGCTACTTTGGAGATCAGCATCGGGA 291
QY 272 CCCCGCTCAAGAGTTGACCGTCTCTTGTATGATGCTGCTTCTGACTTGTGGTTCCT 331
Db 292 CTCACACCCAGAACTTCTGCTGCTTTTGTGACCGGCTCCTCAACTTGTGGGTGCTCT 351
QY 332 CTATCTCTGCAAGAGCAATGCTCTGCAAGAACCAACCAAGATTGATCCGAGAAAGTCT 391
Db 352 CTGTCTACTGCAAGAGCAGCGCTGACACAGTCACTCCGCTTCAACCCAGCAGATGCT 411
QY 392 CCAACCTTCCAGAACTTGAAGAAACCTTGTCTATATACACTAGGATGAGATGAGCAAG 451
Db 412 CCACTCTCTCCACCAATGGGAGACCTTCTCTGCAAGTATGAGAGGACCTTCAACCG 471
QY 452 GAATCTTAGCTATGATACCGTCACTGTCTCCAAATTTGAGATTCATCAACAGACAGTAG 511
Db 472 GCTTCTTTGGCTATGACACCTGACTGTCCAGAGCATTCAGAGTCCCAACAGAGATTCG 531
QY 512 GACTTAGACCCCAAGAACAGAGTATGCTTCACTTATGCAAGATTGATGGCATCTTGG 571
Db 532 GCTTGAATGAGAAAGACCTGTGTACCAACTTCTGTATGGCGAGTTGATGAGCATCATGG 591
QY 572 GTATGGCATCCCATCGCTGCGCGTCAAGATGATCGATACCGTGTGTTGAACAACATGATG 631
Db 592 GCCTGGCTACCTGCTGCTGTCTGCTGATGAGGCCACCAACGATATGCAAGGAGTGGTGC 651
```

```
QY 632 ACCGACACTAGTACTCAAGACTGTCTCGTTTACATGACAGAAATGCCAGAGAGA 691
Db 652 AGAGAGGCGCCCTTCACACAGACCCCGTCTTCAGGGGTCTACTCTCAGAACCAAGAGGCTTCA 711
QY 692 GATGTCTACGCG-----TTGAGAGTATGATTCATTCCTACTACACAGGATCTCTTCACT 745
Db 712 GCGGGGAGCGGTTGCTTCTTTGGGGGTGTGATGACAGCTGTATACAGGGGCGAGATTTACT 771
QY 746 GGGTTCACATCACTGTGACAGATGACTGCAATTCAGTGTGACAGTGTACACATCAGCG 805
Db 772 GGGCGCTGTGTACCCAGAACTCTACTGACAGTTGCGATTGAAGATTCCTCATGCGGG 831
QY 806 GTGTGG---TTGTGATGTGAAGGTGATGTCAAGCTATCTTGGATATCCGTTACCTCCA 862
Db 832 GCCAGGCTCTCGCGCTGTGTCTGAGGGTGTGCGAGGCCATCTGTGACACAGGACACTCTTC 891
QY 863 AGCTGTGCGACCTTGACAGCAGCATTTCTCAACTTCAGACAGTATTTGAAGCCACACAGA 922
Db 892 TGCTTAATGTGCCCGACAGTACATGATGTCTTCTTGACAGCCACAGGGGCCAGAGAG 951
QY 923 ACCAGTACGGTGAAGTTTGAATGATGAGCAACCTTATGATGCTTACAGTGTGCT 982
Db 952 ATGAGATGAGACAGTTTCTGTGAACTGTATGACAGATTCAGAAATCGCCAGCTTGA 1011
QY 983 TTGAGATCAACGGCAAGATGTACCACTGACCCCTCGCTTATACAGCCAGATCAAG 1042
Db 1012 TCATCATCAATGAGTGTGAGGTTCCCTCTCCACCTTCTCTATATCTCAGTAAACAG 1071
QY 1043 GGTTCGACACAGTGAATTCAGAGTGAAGAACATTC-----CAGAAAT 1087
Db 1072 GCTACTGACACCGTGGAGTGAAGCCCACTTACTGTCTCCAGAAAGGCCAGCCCTGT 1131
QY 1088 GGAATCTTGGAGATGTGTTCATTTCTGAGTACTACAGCCTCTTGTGACAGGGCCACAAAC 1147
Db 1132 GAACTCTCGGGAGATGTCTTCTCAGGCTCTATCTATTCGTTACAGACTTGGGCAACAA 1191
QY 1148 TCGTTGGGCTAGCTAAAGCAATCT 1171
Db 1192 GAGTAGGCTTTTGGCACTGCGCCT 1215

RESULT 11
US-09-470-954A-45
; Sequence 45, Application US/09470954A
; Patent No. US20020169304A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: ASPARTYL PROTEASES
; FILE REFERENCE: PC-0008 US
; CURRENT APPLICATION NUMBER: US/09/470,954A
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PERL Program
; SEQ ID NO 45
; LENGTH: 1366
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g1658285
US-09-470-954A-45

Query Match 21.2%; Score 248.4; DB 10; Length 1366;
Best Local Similarity 53.8%; Pred. No. 1.1e-75;
Matches 594; Conservative 0; Mismatches 486; Indels 24; Gaps 3;

QY 92 TTCTCTCTCAAAAGTAAGTCTCTCCGTAAAGCGCTGAAGAAACATGAGCTTCTGAAG 151
Db 112 TGGCCCCGAAAGAAATTTAAGTCTATCCGTAGACCATGAAGAGAAAGGCTTGTGGGGG 171
```

152 ACTTCTTGCAAGAACAGATATGAGCATCAGACAGTACTCCGGCTTCGGTGAAGTTG 211
172 AGTTCCTAGAGACCCAGAAATATGATCTGCTTGGAGTACCGCTTGGTACCTCAGCG 231
212 CTAGGTCGCCATTACCACTACCTGTTAGTCAATTAATCTTTGGGAAGATCTACTCGGAA 271
232 TGACCTACGAGCCCATGAGCTTACATGATGCTGCTTACTTGGTGAATGATCAGCATCGGGA 291
272 CCGCGCCCTCAAGAGTTCAACCGTCTCTTGAATGCTTCCCTGACCTTCGGGTTCCCT 331
292 CTCACCCCAAGAACTTCTGCTGCTCTTCTTGGACACCGGCTCTCCAACTTGTGGGTCCT 351
332 CTATCTACTGCAAGAGCAATGCTGCTGCAAGAACCAACAAAGATTCGATCCGAGAAAGTGT 391
352 CTGCTTAATGCAAGAGCAAGGCTGCAACAGTCACTCCCGCTTCAACCCGAGAGTGT 411
392 CCACCTTCAAGAACTTAGGCAAAACCTTGTCTAATACACTAGCGTACAGGTAGCATGCAAG 451
412 CCACCTACTCAACCAATGAGGAGAGCTTCTCCCTGACAGTATGAGCAGTGGCAGCTCACCG 471
452 GAATCTTAGGCTATGATACCGCTCACTGCTCCAAATGTTGAGCATTTGAACAGACAGTAG 511
472 GCTTCTTGGCTATGACACCTGACTGTCAGAGCATCCAGGTCCCAACGAGAGTTG 531
512 GACTTAGACCAACCAAGAACCAAGGTATGCTTCACTATGAGAAATTCGATGGCATCTTGG 571
532 GCTTAGAGAAATATGAGTACCTGTTACCAACTGCTTATGCGAGATTTGATGATGATCATG 591
572 GTATGAGCAATCCCATCGCTCGGCTGAGAGTACTGATACCTGTTGTTTGAACAATGATGA 631
592 GCTTGGCTTACCTGCTGCTGCTGATGAGGCCACCAACAGCTATGACGGGCTATGGTGT 651
632 ACCGACACTATGATCTAAGAACTTGTCTGCTTATCATGAGCAGGAATGGCCAGAGGA 691
652 AGGAGGGGCGCTTCAACAGCCCGCTTCAAGCTTACTTCAAGCAACGAGGAGCTTCA 711
692 GCATGCTCAGCG-----TTGAGAGCTATTTGATTCATTCCTACTACAGAGATCTTCACT 745
712 GCGGGGAGCGGTTCTTCTTGGGGGTGATGATGACGCTGTGACAGGGGAGATCTTACT 771
746 GGGTTCACGATCACTGTGAGCAGTACTGAGCAATTCATCTGTGAGCAGTGTCACTCAGCG 805
772 GGGGCGCTGTACCCAGGAACCTTACTGAGATTTGGCATTGAAAGATTCCTCATCGGCG 831
806 GTGTGG---TTGTTGATGATGAAGGTGATGATCAAGCTATCTTGAATACCGGTATGCTCA 862
832 GCCAGGCGCTCGGCTGATGTTCTGAGGGTTGCGAGGCCATCGTGGACACAGGCACTCTGC 891
863 AGCTGCTGAGACCTTACAGCGCATTTCTCAATTCAGACAGATTTGAGAGCACAAGA 922
892 TGCTACTGTGTGCCAGCATGATGATGAGTCTTCTGAGGSCCAAGGGGCGCCAGGAGG 951
923 AACAGTACGCTGATGTTGATCATATGATGAGCAACCTTAGTACATGCTTCACTGCTGTCT 982
952 ATGATGATGAGCAAGTTCTCTCGTGAATGTAAGATTCAGAAATCTGCGCCAGCTTGAAGCT 1011
983 TTGAGATCAACGCGCAAGATGTAACCACTGACCCCTCGGCTTAAACAGCCAGATCAAG 1042
1012 TCATCATCAATGATGATGAGTTCCTCTGCGCACTTCTCTATATCTCAATTAACAAG 1071
1043 GGTTCGACCAAGTATGATTCAGATGAGAAACATTTCC-----CAGAAAT 1087
1072 GCTTACTGACCGGTGAGGTGAGGCCAACCTTACTGTCTTCCCAAGAGCGGAGCCCTGT 1131
1088 GGATCTTGGAGATGATGTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1147
1112 GGATCTCTGGGAGATGATCTTCTCTCAGATGCTTACTTATTCGTTACAGATTTGGGCAACAGCA 1191
1148 TCGTTGGGCTAGCTTAAGCAATCT 1171
1192 GAGTAGGCTTTGCGCACTCGCGCT 1215

RESULT 12
US-09-873-319-641
; Sequence 641, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Kulkarni, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Mago, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873.319A
; EARLIER FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 641
; LENGTH: 1366
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 U75272
US-09-873-319-641
Query Match 21.2%; Score 248.4; DB 13; Length 1366;
Best Local Similarity 53.8%; Pred. No. 1.1e-75;
Matches 594; Conservative 0; Mismatches 486; Indels 24; Gaps 3;
92 TTCTCTTCAAGAAAGTATGCTCTCGGTAGAGGCTGAGAGAAATGATGATCTTCTAGAG 151
112 TGCCCTTCAAGAAATTTTAAATGCTTATCGGTGAGACATGAGAGAGAGGCTTGTGGGG 171
152 ACTTCTTGCAAGAACAGATATGAGCATCAGACAGTACTCCGGCTTCGGTGAAGTTG 211
172 AGTTCCTAGAGACCCAGAAATATGATCTGCTTGGAGTACCGCTTGGTACCTCAGCG 231
212 CTAGGTCGCCATTACCACTACCTGTTAGTCAATTAATCTTTGGGAAGATCTACTCGGAA 271
232 TGACCTACGAGCCCATGAGCTTACATGATGCTGCTTACTTGGTGAATGATCAGCATCGGGA 291
272 CCGCGCCCTCAAGAGTTCAACCGTCTCTTGAATGATGATCAATCTTGTGAGTTCCTCT 331
292 CTCACCCCAAGAACTTCTGCTGCTTCTTGAATGATGATCAATCTTGTGAGTTCCTCT 351
332 CTATCTACTGCAAGAGCAATGCTGCTGCAAGAACCAACAAAGATTCGATCCGAGAAAGTGT 391
352 CTGCTTAATGCAAGAGCAAGGCTGCAACAGTCACTCCCGCTTCAACCCGAGAGTGT 411
392 CCACCTTCAAGAACTTAGGCAAAACCTTGTCTAATACACTAGCGTACAGGTAGCATGCAAG 451
412 CCACCTACTCAACCAATGAGGAGAGCTTCTCCCTGACAGTATGAGCAGTGGCATCGGAG 471
452 GAATCTTAGGCTATGATACCGCTCACTGCTCCAAATGTTGAGCATTTGAACAGACAGTAG 511
472 GCTTCTTGGCTATGATACCGCTCACTGCTTCAATGATGATGATGATGATGATGATGATGAT 531
512 GACTTAGACCAACCAAGAACCAAGGTATGCTTCACTATGAGAAATTCGATGGCATCTTGG 571
532 GCTTAGAGAAATATGAGTACCTGTTACCAACTGCTTATGCGCAGTTTATGATGATCATG 591
572 GTATGAGCAATCCCATCGCTCGGCTGAGAGTACTGATACCTGTTTGAACAATGATGA 631
592 GCTTGGCTTACCTGCTGCTGCTGATGAGGCCACCAACAGCTATGACGGGCTATGGTGT 651
632 ACCGACACTATGATCTAAGAACTTGTCTGCTTATCATGAGCAGGAATGGCCAGAGGA 691
652 AGGAGGGGCGCTTCAACAGCCCGCTTCAAGCTTACTTCAAGCAACGAGGAGCTTCA 711
692 GCATGCTCAGCG-----TTGAGAGCTATTTGATTCATTCCTACTACAGAGATCTTCACT 745

Db 712 GCGGGGAGGCGTGTCTTTGGGGGTGTGATAGCAAGCTGTGTACAGGGGCAAGTCTACT 771
Qy 746 GGGTTCCAGTCACTGTGACAGCACTACTGCAATTCATCTGGA CAGTGCACCATGCG 805
Db 772 GGGCGCTGTGACCCAGGAACCTACTGCGAGATGGCATTTGAAGATTCCTCATGGCG 831
Qy 806 GTGTGG---TTGTTCATGTGAAGGTGATGTCAAGTATCTTGGATCCGGTACCTCA 862
Db 832 GCCAGGCTCCGGCTGGTGTCTGAGGGTTGCCAGGCGCATGTGACACAGGCACTCTC 891
Qy 863 AGCTGTGCGGACCTAGCAGCAGCATTTCTCAACATTCGCAAGGCTATTGGACCAAGCA 922
Db 892 TGTCTAATGTGCTCCAGCACTAATGATGCTCTTCTGAGGCGCACAGGGGCCAGAGG 951
Qy 923 ACCAGTACGGTGAAGTTTGACATGATTTGCAACACCTTAGCTACATGCTTACAGTTGTCT 982
Db 952 ATGAGATGAGACAGTTTCTGTGACTGTAAACAGCATTTAGAAATCTCCAGCTTGACCT 1011
Qy 983 TTGAGATCAACGGCAAGATATACCACTGACCCCTCCGCTTATACAGCCAGATCAAG 1042
Db 1012 TCATCATCAATGGTGTGAGATTCCTCTTGCACCTTCTCTATATCTCAGTAAACAG 1071
Qy 1043 GGTTCGACACAGTGTGATTCAGAGTGAAACCATTC-----CAGAAAT 1087
Db 1072 GCTACTGCAACCGTGGAGTGAAGCCCACTACTGTCTCCAGAACGCGCAGCCCTGT 1131
Qy 1088 GGATCTTGGAGATGTGTTCATTCGTGAGTACTACAGCGTCTTTGACAGGGCCAAACACC 1147
Db 1132 GGATCTCGGGGATGTCTTCTCAGGTCCTACTATTCGTCTAGATTGGGCAACACA 1191
Qy 1148 TCGTGGCGCTAGCTAAAGCATCT 1171
Db 1192 GAGTAGGCTTTGGCACCTGCCGCT 1215

RESULT 13
US-09-960-706-984
; Sequence 984, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Mungier, William E.
; TITLE OF INVENTION: Identifying drugs for and diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 984
; LENGTH: 1366
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 U75272
US-09-960-706-984

Query Match 21.2%; Score 248.4; DB 13; Length 1366;
Best Local Similarity 53.8%; Pred. No. 1.1e-75;
Matches 594; Conservative 0; Mismatches 486; Indels 24; Gaps 3;
Qy 92 TTCTCTCTCAAGGTAAGTCTCTCGTTAAGGCGCTGAAGAAACATGACTTCTAAG 151
Db 112 TGCCCTCGAAGAAATTTAACTTATCGTGAACCAATGAAGAGAGGCTTGCGGG 171
Qy 152 ACTTCTTGCAAGAAACAGTATGAGCATGACAGCAAGTACTCCGGGCTTGCGTGAAGTTG 211
Db 172 AGTTCTGAGGACCCACAAATGATGCTGCTTGGAAAGTACCGCTTTGGTGAAGCTGCG 231

Qy 212 CTAGCGTGCCACTTACCACTACCTTGATATGTCATATCTTTGGGAAGATCTACTGGAA 271
Db 232 TGACCTTAGAGCCCAATGAGCTTACATGATGCTGCCATCTTTGGTGAATGATAGATGGA 291
Qy 272 CCCCCTCAAGAGTTTCAACCGTTCTTTGATATGTTGTTCTGTGACTTCTGGGTTCCCT 331
Db 292 CTCACCCCAAGAACTTCCGTGTCTTTTGAACCGGCTCTCTCAACTTGTGGTCCCT 351
Qy 332 CTATCTTAGTCAAGAACAAATGCTGCAAGAACCAACCAAGATTTGATCCGAGAAATGCT 391
Db 352 CTGTCTAGTCCAGAGCCAGGCTTGACACATGCTCCGCTTCAACCCAGGAGTCCGT 411
Qy 392 CCACTTCCAGAACTTAGCCAAACCTTGTCTATATACATACGATAGATGATGATCAAG 451
Db 412 CCACTTCTCCACCAATGGGCAAGACTTCTCCCTGAGATGAGCAATGAGCTTACCG 471
Qy 452 GAATCTTAGGCTATGATACCGTCACTGTCTCCAACTTGTGACATTTCAACAGACATGAG 511
Db 472 GCTTCTTGGCTATGACACCTGACGTCTCAGAGCATTCAGATCCCAACCGAGATTCG 531
Qy 512 GACTTAGCAACCAAGAACAGGTATGTCTTCACTATGCAAGATTTGATGCAATCTTG 571
Db 532 GCTTAGTGAAGATGAGCTTGATCAACATTCGTCTATGCGCAGATTTGATGCAATCATG 591
Qy 572 GATAGCATATCCATGCTCGCTGAGAGTATCCGATACCTGTGTTGACAAATGATGA 631
Db 592 GCTTGCTTACCTGTCTGTCTGTGATGAGGCCCAACAGCTATGACGGGATGTGC 651
Qy 632 ACCGACACCTAGATGATCAAGACTTGTCTCGGTTTACATGACAGAAATGCGCAGAGA 691
Db 652 AGAAGGCGCCCTCACACAGCCCGCTTACAGGCTTACTCAGCAACAGAGGCTCCA 711
Qy 692 GCATGCTCAAGC-----TTGAGCTATTGATTCATCTTACTATACACAGATCTTCACT 745
Db 712 GCGGGGAGCGGTTGTTCTTTGGGGGTGTGATAGCAGCTGTATACAGGGGAGATCTACT 771
Qy 746 GGGTTCCAGTCACTGTGACAGCACTAGTGCATTTCACTGAGCAAGTGTACCATGAGG 805
Db 772 GGGCGCTGTGACCCAGGAACCTTACTGTGCAATTTGAAAGTTCTCTATGCGGCG 831
Qy 806 GTGTGG---TTGTTCATGTGAAGGTGATGTGTCAGACTTCTTGGATCCGGTAGCTCA 862
Db 832 GCCAGGCTTCGGCTGTGTCTTCTGAGGGTTGCCAGGCCATCTGAGACAGGCACTCTTC 891
Qy 863 AGCTGTGCGAAGCTTACAGCAGCAATTTCTCAATTCAGCAAGCTATTGGACCCACAGA 922
Db 892 TGTCTAATGTGACCCAGCATGATAGTGTCTTCTGACAGGCCACAGGGGCCAGAGG 951
Qy 923 ACCAGTACGGTGAAGTTTGACATAGATTGCGCAACCTTAGCTACATGCTTACAGTTGTCT 982
Db 952 ATGAGATGAGACAGTTTCTGTGAACCTGTAAACAGCAATTCAGAAATCTGCCAGCTGACCT 1011
Qy 983 TTGAGATCAACGGCAAGATGTAACCACTGACCCCTCCGCTTATACAGCAGAGATCAAG 1042
Db 1012 TCATCATCAATGTGTGAGTTTCTCTGCACTTCTCTTATATCTCAGTAAACAG 1071
Qy 1043 GGTTCGACACAGTGAATTCACAGAGTGAAGAACCATTC-----CAGAAAT 1087
Db 1072 GCTACTGCAACCGTGGAGTGAAGCCCACTACTGTCTCCAGAACGCGCAGCCCTGT 1131
Qy 1088 GGATCTTGGAGATGTGTTCATTCGTGAGTACTACAGGCTTTTGAACAGGGCCAAACACC 1147
Db 1132 GGATCTCGGGGATGTCTTCTCAGGTCCTACTATTCGTCTACAGCTTGGGCAACACA 1191
Qy 1148 TCGTGGCGCTAGCTAAAGCATCT 1171
Db 1192 GAGTAGGCTTTGGCACCTGCCGCT 1215

RESULT 14
US-10-051-874-53
; Sequence 53, Application US/10051874
; Publication No. US2004000557A1

GENERAL INFORMATION:
 APPLICANT: Padigaru, Muralidhara
 APPLICANT: Alsobrook II, John P
 APPLICANT: Coleman, Steven D
 APPLICANT: Spyrek, Kimberly A
 APPLICANT: Boldog, Ferenc AM
 APPLICANT: Vermet, Corine AM
 APPLICANT: Li, Li
 APPLICANT: Shenoy, Suresh G
 APPLICANT: Casman, Stacie J
 APPLICANT: Guo, Xiaojia Sasha
 APPLICANT: Edinger, Shlomit R
 APPLICANT: MacDougall, John R
 APPLICANT: Malyankar, Uriel M
 APPLICANT: Patutarjan, Meera
 APPLICANT: Shinkets, Richard A
 APPLICANT: Pena, Carol RA
 APPLICANT: Tchernev, Velizar T
 APPLICANT: Zethusen, Bryan D
 APPLICANT: Millet, Isabelle
 APPLICANT: Miller, Charles E
 APPLICANT: Lepley, Denise M
 APPLICANT: Smithson, Glenda
 APPLICANT: Baumgartner, Jason C
 APPLICANT: Herrman, John L
 APPLICANT: Peyman, John A
 APPLICANT: Gorman, Linda
 APPLICANT: Mezes, Peter D
 APPLICANT: Kexuda, Kamesh
 APPLICANT: Taupier Jr, Raymond J
 APPLICANT: Gerlach, Valerie
 APPLICANT: Grose, William M
 APPLICANT: Liu, Xiaohong
 APPLICANT: Ellerman, Karen
 APPLICANT: Rotenberg, Mark
 APPLICANT: Stone, David J
 APPLICANT: Burgess, Catherine E
 TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
 TITLE OF INVENTION: USING THE SAME
 FILE REFERENCE: 21402-245
 CURRENT APPLICATION NUMBER: US/10/051,874
 CURRENT FILING DATE: 2002-09-25
 PRIOR APPLICATION NUMBER: 60/268,595
 PRIOR FILING DATE: 2001-02-14
 PRIOR APPLICATION NUMBER: 60/325,306
 PRIOR FILING DATE: 2001-09-27
 PRIOR APPLICATION NUMBER: 60/262,587
 PRIOR FILING DATE: 2001-01-18
 PRIOR APPLICATION NUMBER: 60/272,409
 PRIOR FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: 60/262,454
 PRIOR FILING DATE: 2001-01-18
 PRIOR APPLICATION NUMBER: 60/276,777
 PRIOR FILING DATE: 2001-03-16
 PRIOR APPLICATION NUMBER: 60/291,672
 PRIOR FILING DATE: 2001-05-17
 PRIOR APPLICATION NUMBER: 60/330,336
 PRIOR FILING DATE: 2001-10-18
 PRIOR APPLICATION NUMBER: 60/265,530
 PRIOR FILING DATE: 2001-01-31
 PRIOR APPLICATION NUMBER: 60/261,376
 PRIOR FILING DATE: 2001-01-16
 NUMBER OF SEQ ID NOS: 269
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 53
 LENGTH: 1270
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-051-874-53
 Query Match 19.7%; Score 230.6; DB 12; Length 1270;
 Best Local Similarity 54.3%; Pred. No. 1.9e-69;
 Matches 514; Conservative 0; Mismatches 424; Indels 9; Gaps 2;

QY 92 TTCTCTTACAAAGTAAGTCTCTCGTAAAGCGCTGAAGAACATGSACTTCTAGAG 151
 Db 69 TGCCCTGAGAAATTTAAGTCTATCGTGAGACCATGAGAGAGAGGCTTGCGGGG 128
 QY 152 ACTTCTGAGAAACAAAGTAAGTGCATCAGACGAACTACTCCGGCTTCGGTGAAGTTG 211
 Db 129 AGTTCTTGAAGAACCAAGTAAGTATCTGCTTGGAAGTATCCGCTTGGATCCTCAGG 188
 QY 212 CTAGCGTGCACCTTACCAACTACCTTGAATGATCACTTGGAGAAATCTACTCGGA 271
 Db 189 TGACTTGAAGGCCATGCGCTTATGATGCTGCTTCTTGGTGAATCAGATTCGGGA 248
 QY 272 CCCCCTCAAGAGTTCAACCGTTCTTTGATGATGTTGTTCTGCTGACTTCTGGGTTCCCT 331
 Db 249 CTCACCCCAAGACTTCTCGTCTTTTGAACCGGCTCTCCCAACTTGTGGGTCCT 308
 QY 332 CTATCTACTGCAAGCAATGCTGCAAGAACACCAAAAGTTGATTCGAGAAAGTGT 391
 Db 309 CTGTCTACTGCAAGCCAGGCTGCAACAGTCACTCCCGCTTCAACCCAGCGAGTGT 368
 QY 392 CCACCTTCCAGAACTTAGGCAAAACCTTGTCTATACCTAGCGTACAGTACAGCAAG 451
 Db 369 CCACCTTCCAGCAAGTGGGAGACCTTCTCTGCAAGTATGCGAGTGCAGCTCACC 428
 QY 452 GAATCTTAGGCTATGATACCGTCACTGCTCCCACTTGTGACATTCACAGACAGTAG 511
 Db 429 GCTTCTTTGGCTATGACACCTGATCTGCAGAGATCCAGTCCCAACAGAGTTG 488
 QY 512 GACTTAGCAACCAAGAACAGGTGATCTTCACTTATGCAAGAAATGATGATGATGCT 571
 Db 489 GCTTAGTGAAGATGAGCTGCTGATCCCACTTGTATGCGAGTTGATGATGATGATG 548
 QY 572 GATAGCATACCCATGCTGCTGCTGATGATGATGATGATGATGATGATGATGATG 631
 Db 549 GCTGCGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 608
 QY 632 ACCGACCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 685
 Db 609 AGAGAGGCGCCCTTCAACGAGCCCTGCTTACGCTTACCTGACCAACAGAGGCTCCA 668
 QY 686 AGAGAGCATGCTACGCTGAGCTATGATGATGATGATGATGATGATGATGATGATG 745
 Db 669 GCGGGGAGCGGTTGCTTTGGGGGTTGATGATGATGATGATGATGATGATGATGATG 728
 QY 746 GGGTTCAGTCACTGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 805
 Db 729 GGGCGCTGTCAACCGAGAACTCTGCTGAGATGATGATGATGATGATGATGATGATG 788
 QY 806 GTGTGG---TTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 862
 Db 789 GCGAGGCTCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848
 QY 863 AGCTGTGCGACCTTACGACGACATTTCAACATTCAGACGATTTGAGAGCACACAGA 922
 Db 849 TGCTACTGTGCCCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 908
 QY 923 ACCAGTACGTTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 982
 Db 909 ATGATGATGACAGTTTCTCGTGAATGATGATGATGATGATGATGATGATGATGATG 968
 QY 983 TTGATGATGACAGGCAAGTATGATGATGATGATGATGATGATGATGATGATGATG 1029
 Db 969 TCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1015

RESULT 15
 US-10-274-639-30
 ; Sequence 30, Application US/10274639
 ; Publication No. US20030232349A1
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE GENOMICS, INC.
 ; APPLICANT: DELEGEANE, Angelo M.; GANDHI, Ameena R.

```

; APPLICANT: HAPALIA, April J.A.; LU, Dyoung Aina M.
; APPLICANT: PATTERSON, Chandra; TRIBOULET, Catherine M.
; APPLICANT: DAS, Depoitiya; KALILICK, Deborah A.
; APPLICANT: NGUYEN, Daniel B.; LEE, Ernestine A.
; APPLICANT: KHAN, Farrah A.; YUE, Henry
; APPLICANT: AU-YOUNG, Janice K.; GRIFEIN, Jennifer A.
; APPLICANT: POLICKY, Jennifer L.; RAMKIDAR, Jayalaxmi
; APPLICANT: YANG, Vunming; THANGAVELU, Kavitha
; APPLICANT: DING, Li; KEARNEY, Liam
; APPLICANT: BAUGHN, Mariah R.; BOROMSKY, Mark L.
; APPLICANT: SANJANMALA, Madhusudan M.; YAO, Monique G.
; APPLICANT: BURFORD, Neil; WALIA, Narinder K.
; APPLICANT: LAL, Preeti G.; LEE, Sally
; APPLICANT: TODD, Stephen; LO, Terence P.
; APPLICANT: TANG, Y. Tom; ELLIOTT, Vicki S.
; APPLICANT: AZIMZAI, Yalda; LU, Yan
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: PI-0167 USA
; CURRENT APPLICATION NUMBER: US/10/274,639
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/22397
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/220,063
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/221,680
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/223,544
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/224,717
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/225,988
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 60/227,568
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 30
; LENGTH: 1173
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030232349A1 7477287CBI
US-10-274-639-30

```

```

Query Match      19.3%; Score 226.8; DB 12; Length 1173;
Best Local Similarity 54.4%; Pred. No. 3.9e-68;
Matches 605; Conservative 0; Mismatches 472; Indels 35; Gaps 6;

```

```

QY      83 TCACCGCATTCCTCTCAAAAGGTAAGTCTCTCCGTAGAGCGCTGAAGAAACATGAGAC 142
DB      50 TCTGCAGATCAATTCGAGGAAAGGCAAGCTATCCGCCAGAGATGAGAGACAGGGTGG 109
QY      143 TTCTAGAAAGACTTCTTGCAAGAAAC-----AGATGGCATCAGCAGAGATGATCCG 196
DB      110 TACTGAGAGCGTTTCTGAGGAGACACCCAAAGGCTGATTCCAATTTGCCAAGTATTTTCA 169
QY      197 GCTTCGTAAGTGTCTAGAGGTGCACTTACCAACTACCTTGATGATCAATACTTTGGGA 256
DB      170 ATATGATATGCTGTGTCTTATGAGCCCTTACCAACTACCTGATTTCTTTTACTTTGGGG 229
QY      257 AATCTTACCTCGAAGACCCGCTCAAGAGTTCAACGTTCTTTGATGATGTTCTCTG 316
DB      230 AGATGACACTGGGACACACCCCAAAATTTTC-CTAGTCTTTGATACGGGTTCTCCA 288
QY      317 A-CTTCTGGGTTCCCTCTATCTTCTGCAAGAGCAATGCTGCAAGAACCCAAAGATTTC 375
DB      289 ATCTGTAGCTGCTGCTCCCACTTACTGCAAGCAAGTCTGCTCCCAATCAACACAGTTTC 348
QY      376 GATCCGAGAAAGCGTCCCACTTCCGAAACTTAGGCAAGACCTTGCTATATACCTAGGGT 435
DB      349 AATCCAGGCTGTCTCCACTTTCAAGAAAGATGAGACAAACTTATGAGACTATCTTATGGG 408

```

```

QY      436 ACAGTAGCATGCAAGGAATCTTAAGCTATGATACCGGTCACTGTCTCCAAATTGGAC 495
DB      409 AGTGGAGCTGATGATGTCTTCTCGGCTATGACATGTAATGCTTCAATACATCGTTGTC 468
QY      496 ATTCAACAGACAGTAGACTTAGACCCCAAGAACAGAGTAGATGCTTCACTATGACAA 555
DB      469 AATTAACAGAGATTGGCTGCTAGTAGAATGAGCCAGAGAACCCCTTTATCTATTCAGTA 528
QY      556 TTGATGATCTCTTGGTATGAGATACCATGCTGCGCTCAGAGTACTGATACCTGTG 615
DB      529 TTGACGGGATCTTGGAATGAGCTACCAACATGACAGAGGGAATTCCTTACAGTA 588
QY      616 TTGACAAATGATGTAACCGACACTAGTAGTCTCAAGACTTGTCTCGGTTTACATGAC 675
DB      589 ATGCAAGGAGATGTGACAGAGACCGACTTACTACCCGCTTCAAGCTTCACTTCACC 648
QY      676 AGGAATG-----CCAGAGACATGCTCAGCCTTGAAGCTATTGATTCATCTAC 726
DB      649 TGCAGCCAAACCGCCAGATATTGTGAGAGCTTCATCTTGGAGTGTGACCCCAACTT 708
QY      727 TACACAGATCTTCTCACTGGGTTCCAGTCACTGTGACAGCACTGGAATT--CACT 783
DB      709 TATTTGTGATCATCTTGTGACCCCTGTCAAGCCGGAACGTGACTGACAGATTGCGATC 768
QY      784 GTGACAGGTCACTCACTCAGCGGTGTGTTGCAATGGAAGTGAATGTCAAGCTATC 843
DB      769 GAGGAATTTGCCATCGTATCAAGGCAAGCACTGCTGTGTGCTGAGGGTTGCCAGGCCATT 828
QY      844 TTGATACCGGTATACGTCCAAAGCTGTGCGAACCTAGACGCACTTCAATTCAGCAA 903
DB      829 GTGATACCGGATACCTTCTGCTGAGATTTCTCAGACATGATGCTTCTTCTGACG 888
QY      904 GCTATTGAGGCCACACAAAGCAATACGATGATTTGACATGATTTGCAACACCTTAGC 963
DB      889 GCAACAGAACCCGACAGGCTCAGAAATGTTGACTTTGTGTCAATGACAGCTATACAG 948
QY      964 TACATGCTTACAGTTGTCTTTGATGATCAACGCGAAGATGTACCCACTGACCCCTCGCC 1023
DB      949 AGCATGCCACCATCACTTTCATCATCGCGGGGCCAGTTTCTCTGCTCTCTTGAA 1008
QY      1024 TATACGACGAGATCAAGGTTCTGACACAGTGAATTCAGAGTGAAGAACATTC--- 1080
DB      1009 TATGTCTTCAATTAACATGCTACTGACAGCTTGAAGTGAAGCCACTGCTGCTCTCC 1068
QY      1081 -----CAGAAATGATCTTTGGAGATGTGTTCAATTCGTAGTACTACAGCGTC 1128
DB      1069 CGCAGTGGGACGCGCTCTGGAATTCGAGGGAATGTCTTCTCAAGGAATATGCTGTGC 1128
QY      1129 TTTGACAGGCGCAACAACTCGTTGGGCTAGC 1160
DB      1129 TATGACATGGCCAAACAGAGGTGGGCTTTTGC 1160

```

Search completed: February 5, 2004, 06:00:52
Job time : 2429 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2004, 01:39:30 ; Search time 2886 Seconds
(without alignments)
9878.431 Million cell updates/sec

Title: US-09-643-755B-1
Perfect score: 1173
Sequence: 1 atgaactccttaagctctt.....ggctagctaaagcaatctga 1173

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrt:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	401.6	34.2	933 13	BQ219820 AGENCOURT
2	387.2	33.0	472 12	BQ938320 IABO15B12
3	358	30.5	631 9	A1892188 mJ84C05.Y
4	354.8	30.2	430 12	BQ937697 IABO05E01

5	337.2	28.7	727 9	A1326975
6	323.8	27.6	399 12	BQ937723
7	312.6	26.6	410 14	CB220875
8	300.8	25.6	1347 11	AK004109
9	300.8	25.6	1408 11	AK053965
10	300.6	25.6	383 12	BQ938086
11	288.6	24.6	623 13	BQ417551
12	271.8	23.2	910 14	CD558204
13	268.8	22.9	826 14	CD516662
14	266.2	22.7	1101 14	CD499194
15	266	22.7	1083 14	CD503881
16	264.4	22.5	1110 14	CD510280
17	264.2	22.5	678 14	BM783853
18	262.6	22.4	1124 14	CD504528
19	262.2	22.4	1157 14	CD496827
20	262	22.3	479 9	AA028632
21	262	22.3	665 12	BM795763
22	262	22.3	691 12	BM817889
23	261.6	22.3	1155 14	CD496442
24	257.6	22.0	605 13	BQ477584
25	257	21.9	1169 14	CD510281
26	256.8	21.9	1122 14	CD509091
27	254.8	21.7	655 12	BM782932
28	254.8	21.7	1224 14	CD505876
29	254.4	21.7	1139 14	CD508825
30	253.4	21.6	443 14	CB220074
31	252	21.5	651 12	BM783855
32	251.4	21.4	1199 14	CD509274
33	251.2	21.4	1197 14	CD505579
34	251	21.4	1140 14	CD497655
35	250.6	21.4	568 12	BQ027751
36	249.6	21.3	505 9	A1479358
37	249	21.2	1134 14	CD498219
38	248.2	21.2	1110 14	CD508053
39	247.8	21.1	307 14	CB221243
40	247.2	21.1	1154 14	CD503882
41	246.4	21.0	297 14	CB220520
42	246	21.0	546 9	A1324867
43	245.2	20.9	1175 14	CD508748
44	243.6	20.8	450 4	BX525462
45	242.4	20.7	517 9	A1385490

ALIGNMENTS

RESULT 1
LOCUS BQ219820 933 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT 7578073 NCI CGAP_Sci Mus musculus cDNA clone
IMAGE:6051528 5', mRNA sequence.
ACCESSION BQ219820
VERSION BQ219820.1 GI:20401220
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Bukaryova, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1305 row: 1 column: 01
High quality sequence stop: 691.

Oy	301	AACATCCAGCAGGCGCATTTGGAGCCACA	CAGAACCA	GTCAGTGA	GTTTGACAT	GACATGC	360
Oy	952	GACAACTTACTACATACCTCTCAAGCTT	TCTTTTGAT	CAACGCCA	GATGTACC	CACTG	1011
Db	361	GACAACTGAGCTACATACCTCCCATCTGT	CTTTTGAG	TAATGGCA	AAATGTACC	CACTG	420
Oy	1012	ACCCCTTCGGCTATTACCGCAGGATGA	GGGTTCTG	CACAGTGA	TTCG	1063	
Db	421	ACCCCTTCGGCTATTACCGCAGGATGA	GGGTTCTG	CACAGTGA	TTCG	472	
RESULT 3							
Al892188							
LOCUS							
DEFINITION	Al892188	631 bp	mRNA	linear	EST 15-MAR-2000		
ACCESSION	U94C05.Y1	Saeres mouse p3NMF19.5	Mus musculus cDNA clone				
VERSION	IMAGE:482792.5	similar to TR:Q28950 Q28950	PREFROCHYOSIN				
KEYWORDS	PRECURSOR ;	mRNA sequence.					
ORGANISM	Al892188						
SOURCE	Al892188.1	GI:5598090					
REFERENCE	EST.						
AUTHORS	Mus musculus	(house mouse)					
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
JOURNAL	Mammalia; Eutheria; Rodentia; Scurionahni; Muride; Murineae; Mus.						
COMMENT	1 (bases 1 to 631)						
	Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rittler ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.						
	The WashU-NCI Mouse EST Project 1999						
	Unpublished						
	Contact: Marra M/WashU-NCI Mouse EST Project 1999						
	Washington University School of Medicine						
	444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA						
	Tel: 314 286 1810						
	Fax: 314 286 1810						
	Email: mousee@watson.wustl.edu						
	This clone is available royalty-free through LMLT ; contact the IMGR Consortium (info@image.llnl.gov) for further information.						
	This read is a RESEQUENCE of a previously sequenced mouse clone This read has been verified (found to hit its original self in the correct orientation)						
	Possible reversed clone: similarity on wrong strand						
	MGI:293536						
	Seq primer: -40RP from Gibco						
	High quality sequence stop: 479.						
FEATURES	Location/Qualifiers						
SOURCE	1..631						
	/organism="Mus musculus"						
	/mol_type="mRNA"						
	/db_xref="taxon:10090"						
	/clone="IMAGE:482792"						
	/dev_stage="19.5 dpc total fetus"						
	/lab_host="DH10B (ampicillin resistant)"						
	/clone_lib="Saeres mouse p3NMF19.5"						
	/note="Vector: pRTT3 (pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCCGACATTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pRTT3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."						
BASE COUNT	147 A	174 C	155 G	155 T			
ORIGIN							
Query Match	30.5%	Score 358;	DB 9;	Length 611;			
Best Local Similarity	73.0%;	Pred. No. 1.1e-92;					
Matches 460;	Conservative	0;	Mismatches 170;	Indels 0;	Gaps 0;		

QY	228	CAACTACCTTGATGATAGTCAATATCTTTGGGAAGATCTACCTCGGAAACCCGCGCTCAAGAGTT	287
Db	1	CAACTACTGATGATGAGAGTACTTTGGAAAGATCTACCTGAGACACCAACCGCAGAGTT	60
QY	288	CACCGTCTCTTTGATATGAGTCTCTCTGACCTTCGGGTTCCCTCTATCTACTGCAAGAG	347
Db	61	CACCGGTGTCTTTTATACAGGCTCTCTCAGAACTCTGGGTACCTCTGTCTACTGCAACAG	120
QY	348	CAATGCGCTGCAAGAACCAACCAAGATTCCGATCCGAGAAATTCGTCCACTTCCAGAACTT	407
Db	121	CAAAGTGTGCGGAACCAACCAACCGTTTGGACCAATCCAACTTCATCCTTCCAGAACTT	180
QY	408	AGGCAAAACCCCTGTGCTCTATACACTACGGTCAAGGTAGCATCAAGAAATCTTAAAGCTATGA	467
Db	181	GAGCAAGACCCCTGTTTGTCCAGTATGTGACTGTGTAATGGAAGGCTTCTGTGCTTACGA	240
QY	468	TACCGTCACTGTCTCCAACTTGTGACATTCAACAGACAGTAGAGACTTAGACCCACAGA	527
Db	241	CAGTGTACAGGTCTCTGATATGTGATGTCCTCAATCAGACTGTGGGCTGAGTACCCACAGA	300
QY	528	ACCAAGGTGATGTCTTCACTCATGCAAGATTCCGATCTCTTGTATGTGCATTAACCATC	587
Db	301	ACCGGGGAGATCTTCACTCACTCACTCTCCATTTGATGGCATCTCGGCGCTGCCTATCTTAC	360
QY	588	GCTGCGCTGACAGATCACTCCGATACCTGTGTTTGAACAATGATGAACCGACACCTAGTAGC	647
Db	361	TTTGTGCTCCAAATATCTCAGTACCCATATTTGACACATATATGAACAGGACCTGTGGTGC	420
QY	648	TCAAGACTTGTCTCGGTTTACATGACAGAGAAATGGCCAGAGAGACATGCTCAACGCTTGG	707
Db	421	CCAAGACCTGTCTTCGGTTTACATGACAGAGAAATGACAGGGGAGCATGTCTCAACTGGG	480
QY	708	AGCTATTGATTCATCTTACTTACACAGAGATCTTCTCACTGGGTTCCAGTACTGTGCACGA	767
Db	481	GCGCATGATCAAGTCCACTTCACTTACAGGCTCACTGCACTGGGTGCTGTGTGACAGTACAGG	540
QY	768	GTACTGGCAATTCACCTGTGACAGTGCACCATCAGCGGTGTGTTGTCATGTGAAG	827
Db	541	ATATTTGCACTTACAGATGACAGATCAACATCAATATGATGAAGTGTGCTTGTCACTTGG	600
QY	828	TGATGTCAAGCTATCTTGGATACCGGTAC	857
Db	601	TGGCTGCTTGGCGGTTCTGACACAGGCAC	630
RESULT 4			
BG937697		430 bp	mRNA linear EST 11-JUN-2001
LOCUS	1AB005E01	Bovine Abomasum cDNA Library	Bos taurus cDNA 5', mRNA
DEFINITION	sequence.		
ACCESSION	BG937697		
VERSION	BG937697.1	GI:14337069	
KEYWORDS	EST.		
SOURCE	Bos taurus		
ORGANISM	Bos taurus (cow)		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos.		
AUTHORS	1 (bases 1 to 430)		
TITLE	Moore,S.S., Hansen,C., Li,C., Fu,A., Meng,Y. and Li,G.		
JOURNAL	cDNA's from bovine abomasum tissue		
COMMENT	Unpublished		
	Contact: Dr. Stephen Moore		
	Beef Genomics Laboratory		
	Dept of AFNS, University of Alberta		
	410 Agr1/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada		
	Tel: 780 492 0169		
	Fax: 780 492 4265		
	Email: stephen.moore@ualberta.ca		
	The sequence best matches gb:BOVCHTMOA (bovine chymosin a (rennin)		
	trna1) in main database at high score of 844.0 and E-value of 0.0		
	PCR Primers		

Db 367 GGGCTGTATTAAGACATCTCAATATTCAGCAAGTCATTTGAGCTGTCGAAGGCCATTA 308
QY 930 CGGTGATTTGACATAGATTGCGACACCTTAGCTACATGACCTGCTTTGAGAT 989
Db 307 TGACCAATTTGACATGCTGCTGAGGCTGACATCATGCCCCAGCTGTTCTTGAGAT 248
QY 990 CAACGGCAAGATGTACCACTGACCCCTCCGCTATACACAGCATCAAGGATCAAGGGCTTG 1049
Db 247 CCATGATGAGGAGTTCCCACTGCAACCTTACCTTATACCAACAGGTCAGGGGCTTCTG 188
QY 1050 CACCACTGATTTCCAGAGTGAAGAACCATTCGCAAAATGGATCTTGGAGATGTTGAT 1109
Db 187 CTCAGATGGCTTC-----AAGCAGGGGCTCCCACTGTGATCTTGGGAGATGTTTCTAT 134
QY 1110 TCGTAGTACTACAGGCTCTTTGACAGGGGCAACACCTGTTGGGCTAGCTTAAGCAT 1169
Db 133 TCGGAGTCTTACAGTGTGTTTGTGACAGACCAACATCGTTGGGCTGGCCAGGCTAT 74
QY 1170 CTGA 1173
Db 73 CTGA 70

RESULT 6
BG937723 399 bp mRNA linear EST 11-JUN-2001
LOCUS 1AB005D06 Bovine Abomasum cDNA Library Bos taurus cDNA 5', mRNA
DEFINITION
ACCESSION BG937723
VERSION BG937723.1 GI:14337095
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bovinae; Bos.
1 (bases 1 to 399)
Moore,S.S., Hansen,C., Li,C., Fu,A., Meng,Y. and Li,G.
CDNA's from bovine abomasum tissue
Unpublished
JOURNAL
REFERENCE Moore,S.S., Hansen,C., Li,C., Fu,A., Meng,Y. and Li,G.
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Dr. Stephen Moore
Beef Genomics Laboratory
Dept of AFNS, University of Alberta
410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
Tel: 780 492 0169
Fax: 780 492 4265
Email: stephen.moore@ualberta.ca
The sequence best matches gb:BFU19786 (Bos primigenius prothymosin
mRNA, complete cds) in main database at high score of 767.0 and
E-value of 0.0
PCR Primers
FORWARD: M13 Forward
BACKWARD: M13 Reverse
Seq primer: T3 primer
High quality sequence stop: 399
POLYA-No.

FEATURES
source Location/Qualifiers
1..399
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/sex="Two males and one female mixed"
/tissue_type="Gastrointestinal tissue (GIT)"
/cell_type="Epithelial"
/dev_stage="Young adult"
/lab_host="X11-BlueMRF"-strain
/clone_lib="Bovine Abomasum cDNA Library"
/note="Organ: Abomasum; Vector: Uni-Z2APK; Site_1: EcoR
I; Site_2: Xho I"

BASE COUNT 89 a 121 c 112 g 77 t
ORIGIN
Query Match 27.6%; Score 323.8; DB 12; Length 399;

Best Local Similarity 88.2%; Pred. No. 7.3e-83;
Matches 352; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 662 CGGTATCATGACACAGAAATGCGCAGAGAGATCTCAAGCTTGAGATATGATCAT 721
Db 1 CGTTTACATGACACAGAAATGCGCAGAGAGATCTCAAGCTTGAGATATGATCAT 60
QY 722 CCTACTACACAGATCTCTTCACTGGGTCCAGTCACTGTGAGAGATAGTACCAATTA 781
Db 61 CCTACTACACAGATCTCTTCACTGGGTCCAGTCACTGTGAGAGATAGTACCAATTA 120
QY 782 CTGTGACAGTGTACCATCACTGAGGCTGTGTTGATGTGATGATGATGATGATGAT 841
Db 121 CTGTGACAGTGTACCATCACTGAGGCTGTGTTGATGTGATGATGATGATGATGAT 180
QY 842 TCTTGATATCCGGTATCGTTCAGAGCTGTGTGACCTAGACAGCACTTCTCAATTCAC 901
Db 181 TCTTGATATCCGGTATCGTTCAGAGCTGTGTGACCTAGACAGCACTTCTCAATTCAC 240
QY 902 AAGCTATTGAGGCGACACAGAACAGTACGGTGTGATGATGATGATGATGATGATGAT 961
Db 241 AAGCTATTGAGGCGACACAGAACAGTACGGTGTGATGATGATGATGATGATGATGAT 300
QY 962 GCTACATGCTTACAGTGTGTTTGAATCAACGCAAGATGTACCACTGACCTCCG 1021
Db 301 GCTACATGCTTACAGTGTGTTTGAATCAACGCAAGATGTACCACTGACCTCCG 360
QY 1022 CCTATACAGCCAGGATCAAGGCTTCTGCACCATGTGAT 1060
Db 361 CCTATACAGCCAGGATCAAGGCTTCTGCACCATGTGAT 399

RESULT 7
CB220875 410 bp mRNA linear EST 10-FEB-2003
LOCUS 1AB027D10 Bos taurus Abomasum #1 library Bos taurus cDNA, mRNA
DEFINITION
ACCESSION CB220875
VERSION CB220875.1 GI:28291389
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bovinae; Bos.
1 (bases 1 to 410)
Hansen,C., Fu,A., Meng,Y., Li,C., Okine,E., Sensen,C.W., Gordon
,P.M.K. and Moore,S.S.
Gene Expression Profiling of the Bovine Gastrointestinal Tract
Unpublished
JOURNAL
REFERENCE Hansen,C., Fu,A., Meng,Y., Li,C., Okine,E., Sensen,C.W., Gordon
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Dr. Stephen Moore
Beef Genomics Laboratory
Dept of AFNS, University of Alberta
410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
Tel: 780 492 0169
Fax: 780 492 4265
Email: stephen.moore@ualberta.ca
Insert Length: 410 Std Error: 0.00
POLYA-No.

FEATURES
source Location/Qualifiers
1..410
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="Smooth muscle"
/cell_type="Epithelial"
/dev_stage="Young adult"
/lab_host="X11-BlueMRF"-strain
/clone_lib="Bos taurus Abomasum #1 library"
/note="Organ: Abomasum; Vector: Uni-Z2APK; Site_1: EcoR
I; Site_2: Xho I"

BASE COUNT 90 a 123 c 120 g 77 t
ORIGIN

Query Match 26.6%; Score 312.6; DB 14; Length 410;
 Best Local Similarity 89.6%; Pred. No. 1.3e-79;
 Matches 336; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 663 GGTTCATGAGACAGGAATGCGCAGAGAGACATGCTCAGCCTTGAGGCTATTGATCCATC 722
 DB 36 GGTTTCAATGACAGAGATGCGCAGAGAGACATGCTCAGCCTTGAGGCTATTGATCCATC 95
 QY 723 CTACTACACAGATCTCTTCACTGCTGCTTCCAGTCACTGTCAGCAGTACTGCAATTCAC 782
 DB 96 CTACTACACAGGCTCCCTGCACTGGGTCGCCGTGACAGTCAAGAGTACTGGCAGTTTAC 155
 QY 783 TGTGACAGTGTCAACATCAACGCGTGTGTTGATGATGGAAGTGTGATCAAGCTAT 842
 DB 156 TGTGACAGTGTCAACATCAACGCGTGTGTTGATGATGGAAGTGTGATCAAGCTAT 215
 QY 843 CTGAGTACCGGAGCGTCAAGCTGTCGAGCCTGACAGGACATCTCAACATTCAGCA 902
 DB 216 CTTGACACGGGACCTTCAAGCTGTGCGGCGCCAGACGACATCTCAACATTCAGCA 275
 QY 903 AGCTATTGAGACACACAGAACAGTACGCGTGAATTTGACATGATTGACCAACTTAG 962
 DB 276 GGCATTTGAGACACACAGAACAGTACGCGTGAATTTGACATGATTGACCAACTTAG 335
 QY 963 CTACATGCTCTACAGTGTCTTTGATCAACGCGAGATGATCCACTGACCCCTTCGC 1022
 DB 336 CTACATGCTCTACAGTGTCTTTGATCAACGCGAGATGATCCACTGACCCCTTCGC 395
 QY 1023 CTATACGACGACGGA 1037
 DB 396 CTATACGACGACGGA 410

RESULT 8
 AK004109
 LOCUS
 DEFINITION Mus musculus 18-day embryo whole body cDNA, RIKEN full-length
 enriched library, clone:1110035E17 product:pepsinogen F, full
 insert sequence.
 ACCESSION AK004109.1 GI:12835163
 VERSION AK004109.1
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 PUBMED 99279253
 MEDLINE 10349636

REFERENCE
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, N., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 PUBMED 20499374
 MEDLINE 11042159

REFERENCE
 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Kitamura, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishu, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yonekura, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 PUBMED 20530913

PUBMED 11076861
 REFERENCE
 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Adachi, J., Fukuda, S.,
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
 Kadota, K., Matsuda, H., Ashburner, M., Batilov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kocichwa, H.,
 Kuehl, P., Lewis, S., Matsuo, Y., Staudli, F., Suzuki, R., Tomita, M.,
 Quackenbush, J., Schriml, L. M., Staudli, F., Suzuki, R., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
 Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bul, C.,
 Gietter, C., Fujita, M., Gariboldi, M., Guelinckx, S., Hill, D.,
 Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
 Marchionni, L., Maehima, J., Mazzarelli, J., Mombarts, P., Nordone, P.,
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
 Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
 Toyo-Oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S.,
 and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 PUBMED 12185660
 MEDLINE 11217851

REFERENCE
 5 The PANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 PUBMED 12185660
 MEDLINE 11217851

REFERENCE
 6 (bases 1 to 1347)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
 Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furumoto, K.,
 Furuno, M., Hasegawa, T., Hara, A., Hayatsu, N., Hiramoto, K.,
 Hirooka, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
 Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
 Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
 Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
 Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
 Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
 Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
 Direct Submission
 Submitted (10-Jul-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 Please visit our web site (http://genome.gsc.riken.go.jp/) for
 further details.
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in Riken.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. First strand cDNA was primed with a primer
 [5' GAGCAGAGAGCGGCCCACTCGATTTTCTTTTCTTTT 3'], cDNA was
 prepared by using tRNAse thermo-activated reverse transcriptase
 and subsequently enriched for full-length by cap-trapper. Second
 strand cDNA was prepared with the primer adapter of sequence [5'
 GAGCAGAGAGAGCGGCCCACTCGATTTTCTTTTCTTTTAAACCCCCCC 3']. cDNA was
 cleaved with XhoI and SclI. Cloning sites, 5' end: SclI; 3' end:
 XhoI. Host: SOLR.
 FEATURES
 Location/Qualifiers
 1. .1347
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM_DB:1110035E17"
 /db_xref="WGI:190662"

/db_xref="taxon:10090"
 /clone="1110035E17"
 /issue_type="whole body"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="18-day embryo"
 40..1203
 /note="unnamed protein product; pepsinogen F
 (MGJ|MGJ:1915935, GB|NM_021453, evidence: BLASTN, 9%,
 match=1210)
 putative"
 /codon_start=1
 /protein_id="BAB23174.1"
 /db_xref="GI:12835164"
 /db_xref="MGJ:1915935"
 /translation="MKMLVGLVALSGLVPIKMKIKSMRENRESOVLDKYLEKY
 PRSAHVLEQRNPATVTEPRNVLIDVYIGISIGTPRPFRLVLTGSSVLTWS
 IYCSPPAHNHRANFPLRSSTFLVSGRPVNAVSGSGFLADVTIVTIDTVVAQA
 FGLSEEPGIFMEYAVFDGILGLVYPNLQGITVFPNLMLOGLIPNLPAFYLSK
 DEKSGMLMGVDPVSYHGLHWVPSKPSYQWLADSIKMGVAVIADCGCGQIMDT
 GTSLLTGRSSIVNIQNLIGAKASGDGEYFLKCDITNLPDVFITGVTVPASAY
 IRKRSNCRSNFBSGMDPSDPEMWVLDVFLRYTFVPRANNRIGLAIPA"
 1332..1337
 /note="putative"
 polyA_site
 1347
 /note="putative"
 polyA_site
 1347
 /note="putative"
 BASE COUNT 296 a 364 c 358 g 329 t
 ORIGIN

Query Match 25.6%; Score 300.8; DB 11; Length 1347;
 Best Local Similarity 58.3%; Pred. No. 7.4e-76;
 Matched 574; Conservative 0; Mismatches 392; Indels 18; Gaps 2;

208 GTTGTACGCGCCACTTACCACTACCTTGTATGATCACTATTGGGAAGTCTACCTC 267
 220 GTAACTTATGAGCTATGAGAACTTACCTGCTGTCTTACATTTGGATCAGCAATT 279
 268 GGAACCCGCGCTCAGAGTTCACCGTCTCTTGTATGATGCTTCTGACCTTCTGGGT 327
 280 GGACGCGCCCTCAGAGTTCACCGTCTCTTGTATGATGCTTCTGACCTTCTGGGT 339
 328 CCTCTATCTACTGCAAGACCAATCCCTGCAAGAACCAAGATTCGATCCGAGAAAG 387
 340 CCATCATCTATTTGCTCAGCGCCAGCTGCGCTGCAACCAAGATTCGATCCGAGAAAG 399
 388 TCGTCCACCTTCCAGAACTTACGAAACCCCTTGTATGATGCTTCTGACCTTCTGGGT 447
 400 TCTTCACTTCTCTGCTCAGCGCCAGCTGTAATGTTGCTTATGCTTCTGAGAGATG 459
 448 CAAGGATCTTACGCTATGATACGCTGCTGCAACATTTGAGACATTTCAAGACA 507
 460 TCCGATTTCTGCTATGACCTGTCAGAGATTGGGACCTTACGCTTGGCCCAAGGCC 519
 508 GTAGGACTTACGACCAAGAACCAAGATGATGCTTCTGACCTTACGAGAAATTCGATG 567
 520 TTGGCTGAGCTGAGAACCAAGAACCTTCTGAGAAATGATGCTTCTGATGATG 579
 568 CTGCTATGACCTTACCACTTCTGCTGCTGCTTACGAGAAATTCGATGATG 627
 580 CTGGGCTGAGTATCCCAACCTTCTGCTGCTTACGAGAAATTCGATGATG 639
 628 ATGACCGACCTTACGCTTACGCTTCTGCTGCTTACGAGAAATTCGATGATG 684
 640 TGGCTACAGGCTTATCCCTGAGAAATCTTCTGCTTCTTACGAGAAATTCGATGATG 699
 685 CAGAGACATGCTCAGCTTGGAGCTATGATGCTTCTTACGAGAAATTCGATGATG 744
 700 AAGGCGAGATGCTGATGCTGAGAGTATGCTTCTTACGAGAAATTCGATGATG 759
 745 TGGGTTCCAGTACCTGAGAGCTTACGAGAAATTCGATGATGCTTCTTACGAGAAATTCGATG 804
 760 TGGGTTCCAGTACCTGAGAGCTTACGAGAAATTCGATGATGCTTCTTACGAGAAATTCGATG 819
 805 GGTGTGTTGTTGATGATGAGTGTCAAGCTATCTTGGATACCGGTACGTCAG 864

820 GGGGAGCTATGCTGCTGATGATGCTGCTGCAAGGATTTATGACACAGGACCTCTTG 879
 865 CTGCTGCGACCTTACGAGCAATTTCTCAACTTACGAGCAAGCTATTTGAGCCACAGAAC 924
 880 CTGACCGGCGCCCGCAAGCTTCAATGTTTACATCCAGAAATTTGTTGCGCAAGGCTCT 939
 925 CAGTACGCTGAGTTTGAATGATTTGAGCAACCTTACCTACATGCTTACATGTTCTT 984
 940 GTGACGCGAGTACTTCTTCAAGTGTACACCTTACATCCACCTTCTGATTTCTTCT 999
 985 GAGATCAACGCGCAAGATGATACCACTGACCCCTTCCGCTTATACGAGCAGATCAAGG 1044
 1000 ACCATGCGAGGTTACTTACCTTACCAAGTGCAGGCTGCTTACATCGAAAGAGTGTCA 1059
 1045 TTCTGACCACTGATTTCAAGTGTGAGAAC-----CATTTCCAGAAATGG 1089
 1060 CACATTCGAGAGCACTTTTGAAGAGGCAATGATGACCATGACCTTCAAGATGTTG 1119
 1090 ATCTTGGAGATGTTTCAATTCGTGATGATGACAGCTTGTGACAGGCGCAACACCTC 1149
 1120 GTGCTGGGAGATGCTTCTTCAAGCTGATTTTACCGGTGTTGATGCGGCAATATACAG 1179
 1150 GTTGGGCTAGCTTAAAGCAATCTGA 1173
 1180 ATTGCTGCTGCTCTCTGCTGATGA 1203

RESULT 9
 AK053965
 LOCUS
 DEFINITION
 Mus musculus 2 days pregnant adult female oviduct cDNA, RIKEN full-length enriched library, clone: E23008A19 product: pepsinogen F, full insert sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 AK053965
 AK053965.1 GI:26343894
 HTC; CAP trapper.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komu, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Komu, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, N., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

REFERENCE
 AUTHORS
 4
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Komu, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R.,

Db 1062 ACCATCGGACGTGTTACCTACCCAGTCCAGCCAGTGGCTTACATCCGAAGATGCTCA 1121
 QY 1045 TTCTGCACACAGTGAATTCAGAGTGAAC-----CATTCGCAAGAAATGG 1089
 Db 1122 CACAAATTCAGAGCAACTTTGAGAGGCAATGATGACCATCAGACCTGAGATGTGG 1181
 QY 1090 ATCTTGAGAGATGTTGTTTCAATTCGTGAGTACTACAGACGTTCTTTGACAGGGCCAAACCTC 1149
 Db 1182 GTGCTGGGGAGATGTTCTCTGAGGCTGATTTTCACTGTTTGAATCGGGCAAAATGAACG 1241
 QY 1150 GTTGGGCTAGCTAAAGCATCTGA 1173
 Db 1242 ATTGCTGCTCTCTGCTGATGA 1265

RESULT 10
 BG938086 383 bp mRNA linear EST 11-JUN-2001
 LOCUS LAB011A08 Bovine Abomasum cDNA Library Bos taurus cDNA 5', mRNA
 DEFINITION
 ACCESSION BG938086
 VERSION BG938086.1 GI:14337458
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 383)
 Moore,S.S., Hansen,C., Li,C., Fu,A., Meng,Y. and Li,G.
 CDNA's from bovine abomasum tissue
 JOURNAL Unpublished
 COMMENT Contact: Dr. Stephen Moore
 Beef Genomics Laboratory
 Dept of AFNS, University of Alberta
 410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
 Tel: 780 492 0169
 Fax: 780 492 4265
 Email: stephen.moore@ualberta.ca
 The sequence best matches gb:BF019786 (Bos primigenius prothymosin
 mRNA, complete cds) in main database at high score of 735.0 and
 E-value of 0.0
 PCR Primers
 FORWARD: M13 Forward
 BACKWARD: M13 Reverse
 Seq primer: T3 primer
 High quality sequence stop: 383
 POLYA=No.

FEATURES
 source Location/Qualifiers
 1..383
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /sex="Two males and one female mixed"
 /tissue_type="Gastrointestinal tissue (GIT)"
 /cell_type="Epithelial"
 /dev_stage="Young adult"
 /lab_host="XL1-BlueMRF-strain"
 /clone_id="Bovine Abomasum cDNA Library"
 /note="Organ: Abomasum; Vector: Uni-22APKX; Site_1: EcoR
 I; Site_2: Xho I"

BASE COUNT 83 a 111 c 113 g 76 t
 ORIGIN
 Query Match 25.6%; Score 300.6; DB 12; Length 383;
 Best Local Similarity 88.3%; Pred. No. 4e-76;
 Matches 338; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

QY 559 AGTACTCATCTCTGTTTGACACATGATTAACGACCTAGTACTCAAGCTTGT 658
 Db 1 AGTACTCATCTCTGTTTGACACATGATTAACGACCTAGTACTCAAGCTTGT 60
 QY 659 TCTCGTTTACATGACAGATGGCCAGAGAGATGCTCAGCTTGAGCTATTGATC 718

Db 61 TCTCGTTTACATGACAGAGATGGCCAGAGACATGCTACCCCTGGGGCCATCGACC 120
 QY 719 CATCTCACTACAGAGATCTCTTCACTGGGTTCCAGTACCTGTGAGAGATCGCAAT 778
 Db 121 GGTCTTACTACACAGGGGTCCTTGACATGGGTCGCCGTGACAGTGCAGATCGACGT 180
 QY 779 TCACGTGGAAGAGTGCACCATCAGCGGTGTGTTGTCATGTGAAGTGGATGTCAG 838
 Db 181 TCACGTGGAAGAGTGCACCATCAGCGGTGTGTTGTCATGTGAAGTGGATGTCAG 240
 QY 839 CTATCTTGATATACCGGTACGTCCAGCTGTGCGACCTAGACGCGACATTCACATT 898
 Db 241 CCATCTCTGACACAGGACCTCCAGAGCTGTGCGGCCACAGCAGACGACATCTCAACATCC 300
 QY 899 AGCAAGCTATGAGACCCACACAGAACCATGATCGGTGATTTGACAT-AGATTGGACAAAC 957
 Db 301 AGCAGGCCATTTGAGCCACACAGAACCATGATCGGTGATTTGACATTTGACTGCGACAAAC 360
 QY 958 CTTAGCTACATGCTCCTACAGTTGT 980
 Db 361 CTGAGCTACATGCTCCTACAGTTGT 383

RESULT 11
 B0417551/c 623 bp mRNA linear EST 23-MAY-2002
 LOCUS IK48c09.x1 Kaestner ngm3 wt Mus musculus cDNA 3' similar to
 DEFINITION SW:CHYM SHEEP P18276 PROCHYMOSIN PRECURSOR, B-FORM ;, mRNA
 sequence.
 ACCESSION B0417551
 VERSION B0417551.1 GI:21122752
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 623)
 Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
 Lemshka,I., Scearce,M., Brestelli,D., Gradwohl,G., Clifton,S.,
 Hillier,L., Mays,M., Page,D., Wylie,T., Martin,J., Blaisdell,A.,
 Schmitt,A., Theising,B., Ritter,B., Ronko,I., Bennett,J., Cardenas
 ,M., Gibbons,M., McCann,R., Cole,R., Tsagarisvilli,R., Williams,T.,
 Jackson,Y. and Bowers,Y.
 Endocrine Pancreas Consortium
 Unpublished
 JOURNAL Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 COMMENT Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@iobp.harvard.edu
 Pancreas was obtained from Gerard Gradwohl (PNAS 97 1607-1611,
 2000) library was constructed by Catherine Lee DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Marie Scearce
 (mscearce@mail.med.upenn.edu)
 Seq primer: -40UP from Gibco
 High quality sequence stop: 382.

FEATURES
 source Location/Qualifiers
 1..623
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129/SV x CDI"
 /db_xref="taxon:10090"
 /dev_stage="p.c. 14.5"
 /lab_host="E. coli-DH12S (GIBCO)"
 /clone_id="Kaestner ngm3 wt"
 /note="Organ: pancreas; Vector: pSPORT1 (GIBCO); Site_1:
 Not I; Site_2: Sal I; The library was prepared by
 Catherine S. Lee and has not been published. The pancreas

was obtained from Gerard Gradwohl (PNAS 97 p1607-1611, 2000). The cDNAs were prepared with an oligo containing a NotI site, and SalI linkers were added to the ends. The inserts were cut with NotI before being cloned into the NotI-SalI sites in the vectors. This is one of two libraries, ngm3 wt and ngm3 -/- . The wt library is in pSPORI, T7 promoter is 5'."

BASE COUNT 152 a 163 c 155 g 153 t
ORIGIN

Query Match 24.6%; Score 288.6; DB 13; Length 623;
Best Local Similarity 71.8%; Pred. No. 1.6e-72;
Matches 394; Conservative 0; Mismatches 149; Indels 6; Gaps 1;

QY 625 ATGATGACCGACCTAGTACTGCAAGCTTGTCTCGTTTACATGACAGGAATGCC 684
DB 623 ATGATGACCGACCTAGTACTGCAAGCTTGTCTCGTTTACATGACAGGAATGAG 564
QY 685 CAGGAGAGCATGTCCGCTTGGAGCTATGATCCATCCATACACAGATCTCTTAC 744
DB 563 CAGGAGAGCATGTCCGCTTGGAGCTATGATCCATCCATACACAGATCTCTTAC 504
QY 745 TGGGTTCCAGTCACTGTGACAGTACTGCAATTCATCTGGACAGTGTCAACATCAGC 804
DB 503 TGGGTTCCAGTCACTGTGACAGTACTGCAATTCATCTGGACAGTGTCAACATC 444
QY 805 GGTGTGTTGTTGCAATGTGAAGTGTGCAAGCTATCTTGGATACCGGTACGTCAG 864
DB 443 GGTGAAGTGTGTTGCAATGTGAAGTGTGCAAGCTATCTTGGATACCGGTACG 384
QY 865 CTGTGTGGACCTTACGACGCACTTCTCAATTCAGCAAGCTATTGGAGCACACGAAC 924
DB 383 TTGACGGGCTGTGTGAGACATCTCAATTCAGCAAGCTATTGGAGCACACGAAC 324
QY 925 CAGTACGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 984
DB 323 CAGTACGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 264
QY 985 GAGATCAACGCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1044
DB 263 GAGATCAACGCGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 204
QY 1045 TTCTGACACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1104
DB 203 TTCTGACACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 150
QY 1105 TTGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1164
DB 149 TTGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 90
QY 1165 GCAATCTGA 1173
DB 89 GCTATTTGA 81

RESULT 12
CD558204 910 bp mRNA linear EST 11-JUN-2003
LOCUS AGENCOURT 14477618 NIH MGC 181 Homo sapiens cDNA clone
DEFINITION IMAGE:30355411 5', mRNA sequence.
ACCESSION CD558204
VERSION CD558204.1 GI:31584272
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 910)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics

National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: NDAM475 row: h column: 12
High quality sequence start: 20
High quality sequence stop: 617.
Location/Qualifiers

FEATURES

source

1..910
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30355411"
/tissue_type="White Matter"
/dev_stage="Unknown"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/note="Vector: PCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV (destroyed); Library is oligo-dT primed and directionally cloned (BcoRV site is destroyed upon cloning). Average insert size 1.42 kb. Library was constructed by (Invitrogen). Note: this is a NIH-MGC Library."

BASE COUNT 184 a 305 c 231 g 190 t
ORIGIN

Query Match 23.2%; Score 271.8; DB 14; Length 910;
Best Local Similarity 62.8%; Pred. No. 1.6e-67;
Matches 439; Conservative 0; Mismatches 257; Indels 3; Gaps 1;

QY 245 AATACCTTGGGAAGATCTACTCTCGAAGCCCGCTCAAGATTCACCTTCTTGATA 304
DB 30 AATACCTTGGGAAGATCTACTCTCGAAGCCCGCTCAAGATTCACCTTCTTGATA 89
QY 305 CTGTGTTCTTGTGATCTTGGGTTCCCTTATCTACTGCAAGAGCAATGCTTGCAGAAC 364
DB 90 CTGTGTTCTTGTGATCTTGGGTTCCCTTATCTACTGCTCCAGTTGCTTGCAGAAC 149
QY 365 ACCAAGATTCGATCCGAGAAAGTGTCCACCTTCAGAACTTGAAGCAACCTTGTCTA 424
DB 150 ACCAAGATTCGATCCGAGAAAGTGTCCACCTTCAGAACTTGAAGCAACCTTGTCTA 209
QY 425 TACACTAGGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 484
DB 210 TACACTAGGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 269
QY 485 ACATGTGACATTCACAGAGTAGAGATTAGACCCAGAGACAGAGATGTCTCA 544
DB 270 ACATGTGACATTCACAGAGTAGAGATTAGACCCAGAGACAGAGATGTCTCA 329
QY 545 CTTATGAGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 604
DB 330 CTTATGAGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 389
QY 605 CGATTCCTGTTGTCACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 664
DB 390 CGATTCCTGTTGTCACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 449
QY 665 TTTACATGACAGGAATGGCCAG--GAGAGCATGTCAAGCTTGAAGCTATTCAT 721
DB 450 TTTACATGACAGGAATGGCCAG--GAGAGCATGTCAAGCTTGAAGCTATTCAT 509
QY 722 CTTACTACACAGATCTTCACTGGGTTCACTGTGACAGTGTGACAGTGTGACATTC 781
DB 510 CTTACTACACAGATCTTCACTGGGTTCACTGTGACAGTGTGACAGTGTGACATTC 569
QY 782 CTGTGACAGTGTCAACATCAGGCTGTGTTTTCATGTGAAGTGTGATGATCAAGCTA 841

Db	570	CCGTGACAGCATCACATGAACGGAGAGGCCATCGCTGGCGTGAAGGGCTGCCAGGCCA	629
Qy	842	TCTTGGAATACCGGGTACTGCITCAAAAGTGGTCGGAAGCCTTAGACGAGCAATTCTCAACATTAGC	901
Db	630	TTGTGTGACACCGGGGACCTCTCTGCTGTAACCGGACCACAAGCCCCCATTCCAAATCCAGA	689
Qy	902	AAGCTATTGTGAGCCACACAGAACCAAGTACGGTGAATTGG	940
Db	690	GCGACATTCGAGGCCAGCGAGAACTCAGATGGCGGACATGG	728
 RESULT_13 CD516662			
LOCUS	CD516662	826 bp	mRNA linear EST 06-JUN-2003
DEFINITION	AGNOCOURT 14371339 NIH MGC 181 Homo sapiens CD51 clone		
	IMAGE:30407989 5', mRNA sequence.		
ACCESSION	CD516662		
VERSION	CD516662.1	GI:31448380	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 826) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Daniela S. Gerhard, Ph.D.		
JOURNAL			
COMMENT			

National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: ccgabs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINTL at:
<http://image.llnl.gov>
Plate: NDAM490 row: d column: 14
High quality sequence stop: 695.

FEATURES	Location/Qualifiers
source	1. .826

```

/organism="Homo sapiens"
/mol type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:30407989"
/issue_type="White Matter"
/dev stage="unknown"
/lab host="DH10B-Ton A ( T1 and T5 phage resistances )"
/clone_1lb="NIH_MGC_181"
/notes="vector: pCMV-SPORT6.1; site_1: NotI; site_2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(invitrogen). Note: this is a NIH_MGC Library."

```

Query Match	22.9%	Score 268.8	DB 14	Length 826
Best Local Similarity	62.6%	Pred. No. 1,le-66		
Matches	436	Conservative	0	Mismatches 257; Indels 3; Gaps 1

QY	219	GCCACTTACCAACTACCTTGATAGTCATATCTTTGGAAAGATCTACCTCGAACC	CCGCC	278
Db	70	GCTCTTGAGAACTACCTCGATATGGAAGTACTCTCGCACTACCGACATCGGAAC	CTCTG	129
QY	279	TCAAGATTCACCGTCTCTTGATAGTACGTCCCTGTGACTCTGGGTCTCCCTATCTA		338
Db	130	CCAGATTTCACCGTCTCTTGACACCGGCTCTTCCAACTGTGGGTGCTCCCTCA	GTCTA	189
QY	339	CTGCAAGAGCAATGCTCTGCAAGAACCAACCAAGATTCATCCGAGAAAGTGTCA	ACTTT	398

Db	190	CTGCTCCAGTCTTGCTGCTGACCAACCAACAAACCGCTTAAACCTGTAGATTTCTTCCACCTA	249
OY	399	CCAGAACTTAGGCAAAACCTTGTTCTATACACTACGGATACAGGTAGCATGCAAGAACTTT	458
Db	250	CCAGTCCACCGACGACAGCTCTCCATCACTACCGGCAACCGGACGATGACAGGATCCT	309
OY	459	AGGCTATGATACCGTCACTGTCTCCAACTGTGACATTGAACGACAGTGGACTTGG	518
Db	310	CGGATACGACACTGTCCAGGTGGAGGATCTTCGACACCAATCAGATCTTTCGGCTTGAG	369
OY	519	CACCCAAAGAACGAGTGTATGTCTTCAACCTATGACAGAAATTGCATGSCATCTTGATGAC	578
Db	370	CGAAGCGAACTGTGCTCTCTCTGTATTAATGTCTCCCTTCGATGCGATCTTGGGCTGCG	429
OY	579	ATACCCATTCGCTCGCGTCAGAGTACTCGATACCTGTGTGTCACACATGATGAACCGACA	638
Db	430	CTACCCCGAGATTTCTCTCTCCGGGGCCACACCGGTCTTTGACAACTGTGAAACCAAGG	489
OY	639	CTTAGTACCTCAAGACTTGTCTCTGGTTTAACTGACGAGAAATGGCCAG--GAGAGCAT	695
Db	490	CTGGTGTTCAGGACCTCTTCTGTCTACTCCTCAGCCCGCAATGACAGAGTGGACGCT	549
OY	696	GCTACGCGTTGGAGCTATTTGATTCATCCTCACTACACAGAGATCTCTTCACTGGGTTCCAGT	755
Db	550	GGTATCTTTTGTGGCATTTAGCTCTTTCTTACTACCTGGAAGTCTGAATCTGGTGCTGT	609
OY	756	CACGTGTGACGACTACTGGCAATTCACTGTGGACAGTGTCAACCATCAGCGGTGTGTTGT	815
Db	610	TACGTTGAGGGTTACTGTGGAGATTCACCGTGTGACGATCAACATGAAACGAGAGGCCAT	669
OY	816	TGCATGTGAAAGTGGATGTCAAGCTATCTTGGATACCGGTACGTTCAAGCTGTGAGACC	875
Db	670	CGCCTGGCGTGGAGGCTGTGCGAGGCCATTTGTGACACCGGCACTCTGTCTGACCGGCC	729
OY	876	TAGACGCGAATTTCTCAACTTCGACGAAGCATTTGG	911
Db	730	AACGAGCCCATTTGCCAAACATTCAGAGCGACATCGG	765

RESULT 14	
CD499194	
LOCUS	
DEFINITION	CD499194 1101 bp mRNA linear EST 12-JUN-2003
ACCESSION	CD38-A02.x1d-t SHC-CD4 Gasterosteus aculeatus cDNA clone
VERSION	CD38-A02.5', mRNA sequence.
KEYWORDS	CD499194 CD499194 GI:31426225
SOURCE	EST.
ORGANISM	Gasterosteus aculeatus (three spined stickleback)
	Gasterosteus aculeatus

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 1101)	Kingsley,D.M., Petchel,C., Balabhadra,S., Grimwood,T., Dickson,M., Schmutz,J. and Myers,R.M.	Expressed sequence tags from <i>Gasterosteus aculeatus</i>	Unpublished
	Contact: Kingsley, DM		HHMI and Department of Developmental Biology	
	Stanford University School of Medicine			
	Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA			
	Tel: 650 725 5954			
	Fax: 650 725 7739			
	Email: Kingsley@cmgm.stanford.edu			
	Plate: 38			
	High quality sequence, stop: 904.			
	Location/Qualifiers			

```
FEATURES
source
location/Qualifiers
1..1101
/organism="Gasterosteus aculeatus"
/mol_type="rRNA"
/strain="Salinas river, CA"
/db_xref="taxon:69293"
```

```

/clone="CDA38-A02"
/sex="mixed male and female"
/tissue_type="heads and internal organs combined"
/dev_stage="adult"
/clone_lib="SHGC-CDA"
/notes="Vector: lambda ZAP Express/pBK-CMV; Site 1: EcoRI
(5' adaptor); Site 2: XhoI (3' linker primer). The mixed
organ cDNA library was generated using the ZAP-cDNA method
by Stratagene. First strand cDNA synthesis was primed with
a 50 bp linker primer containing an oligo dt sequence
preceded by a synthetic XhoI site. 5 prime adaptors were
used containing an EcoRI cohesive end. The finished cDNAs
were inserted in to the ZAP express vector
unidirectionally in the sense orientation with respect to
the lacZ promoter of pBK-CMV. An amplified library was
prepared from approximately 3 million primary clones in
the lambda ZAP Express vector. In vivo excision was then
used to generate individual pBK-CMV phagemid clones for
EST sequencing."
BASE COUNT      268 a      313 c      273 g      243 t      4 others
ORIGIN

```

```

Query Match      22.7%; Score 266.2; DB 14; Length 1101;
Best Local Similarity 56.7%; Pred. No. 7.6e-66;
Matches 548; Conservative 0; Mismatches 412; Indels 7; Gaps 3;

```

```

210 TGGTACGCGCCACTTACCACTTGTATGATCACTTTGGAGAGATCACTCGG 269
270 AACCCGCGCTCAGAGATTCACCTTCTTTGATCTGTCTTGTACCTTCGGTTC 329
141 TAACTCTCTCAGCTCTTCACTGCTCATCTTTGACACCGGCTCTCCAACTG 200
330 CTCTATCTACTGCAAGACATCTCTGCAAGAACCAACCAAGATTCGATCCGAA 389
201 TTCACTCTACTGTCTCAGCAGAGCTTGCACCAACCAATTCACCCAGGAGTC 260
390 GTCCACCTTCCAGACTTGAAGCAACCTTGTCTTATACACTGACGTAAGATGCA 449
261 CAGCAGCTTTCATGAGGAGGAGCAGACTGTCTCATTCAGTACGCGACCGGAG 320
450 AGGAATCTTAGGCTATGATACCTCTACTGTCTCCACATTTGTGACATTTAA 509
321 TGGACGTCTGCAAGATGCAATTTGAGGAGGAGATCAAGGTCCAAACAGGTGT 380
510 AGGACTTAGACCCCAAGACAGATGATGCTTCACTATGAGAAATTCAGTGGCAT 569
381 TGGGATCAGGAGACAGA--ACCTCTTTCATGCTCAGATGACGCCAGTGGCAT 437
570 TGGTATGGAATCCATCGCTCGCTCAGATGACTCGATACCTGTGTTTGAACA 629
438 GGGGTTGGCTTCAGAGCATGCTCTGCAATGTCGTCTTTGACAAACATGTT 497
630 GAACGACACCTTAGTACCTAAGATTTGTTCTGGTTTATAGACAGAAATGCG 689
498 CAGACGAGGAGCTTTGTTCCAGTCTTGTCTCTGCTTACCTGAGGCGCAAT 557
690 GAGCA--TGCTACGCTTGGAGCTATGATCATCTTACATACAGAGATCTTCA 746
558 GGGCAGTGAAGGTGGCTTCGTTGATGACAGTACATTAACCGGTAAATCA 617
747 GGTTCAGTCACTGTGACAGATGAGCAATTCACCTGAGACAGTCCATCAGCG 806
618 GATCCCTCTGACCTCTGCAACCTAATGCGAGATCCAAATGACAGTTCACAT 677
807 TGTGTTGTTGATGATGAGAGGTGATGTCAGCTATCTTGAATACCGGTA 866
678 AAGAGCTGTGCTGCTGACCTGTGCTGCGCAGCATATGACATGATCTCCAG 737
867 GGTCCGACCTAGACAGGACATTTCAACATTCAGCAAGCATATGAGCCACAGA 926
738 TGTGGCCCAACCAATGACATCAACAATGAATTTCTGTGTTGAGCCACAGCCA 797

```

```

QY      927 GTACGGTGAGTTTGACATGATTTGCGACAACTTACTCATGCTTACAGTTCTTTGA 986
DB      798 GATGAGAGAGGCTTACAGTGAATCGCCAAACATCCAGGACATGCTGAAGTACCTTAC 857
QY      987 GATCAACGGCAAGATGTATGCCACTGACCCCTCCGCTTATACAGGACGATCAAGGTT 1046
DB      858 TCTCAACGACAGAGGCTTATCCATCTCTGATCTGCTTACATGACAGCTCTTACGG 917
QY      1047 CTGACCACTGATTCAGAGTGAAGAACATTTCCAGAAATGATTTGGAGATGTT 1106
DB      918 CTGACAGCACTGGCTT-TGGCCAGGAGTGGCTCCACAGCTCTGGATCTGGAGATGTTCTT 976
QY      1107 CATTCGTGATCTACATAGGCTCTTTTACAGGGCCAAACCTTGTGGCTACTTAAAG 1166
DB      977 CATCANGACAGTCTTACATCTTCCATACGAGATCAACAAAGTGTGCTGGCCCATC 1036
QY      1167 AATCTGA 1173
DB      1037 TGTCTTA 1043

```

```

RESULT 15
CD503881      1083 bp      mRNA      linear      EST 12-JUN-2003
LOCUS      CDA64-G12.x1d-t SHGC-CDA Gasterosteus aculeatus cDNA clone
DEFINITION      CDA64-G12 5', mRNA sequence.
ACCESSION      CD503881
VERSION      CD503881.1 GI:31434449
KEYWORDS      EST.
SOURCE      Gasterosteus aculeatus (three spined stickleback)
ORGANISM      Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 1083)
REFERENCE
AUTHORS      Kingsley,D.M., Peichel,C., Balabhadra,S., Grimwood,J., Dickson,M.,
Schultz,J. and Myers,R.M.
TITLE      Expressed sequence tags from Gasterosteus aculeatus
JOURNAL      Unpublished
COMMENT      Contact: Kingsley, DM
HMTI and Department of Developmental Biology
Stanford University School of Medicine
Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
Tel: 650 725 5954
Fax: 650 725 7739
Email: kingsley@cwmg.stanford.edu
Plate: 64
High quality sequence stop: 867.
Location/Qualifiers
1..1083
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Salinas river, CA"
/db_xref="taxon:69293"
/clone="CDA64-G12"
/sex="mixed male and female"
/tissue_type="heads and internal organs combined"
/dev_stage="adult"

```

FEATURES

```

source
/notes="Vector: lambda ZAP Express/pBK-CMV; Site 1: EcoRI
(5' adaptor); Site 2: XhoI (3' linker primer). The mixed
organ cDNA library was generated using the ZAP-cDNA method
by Stratagene. First strand cDNA synthesis was primed with
a 50 bp linker primer containing an oligo dt sequence
preceded by a synthetic XhoI site. 5 prime adaptors were
used containing an EcoRI cohesive end. The finished cDNAs
were inserted in to the ZAP express vector
unidirectionally in the sense orientation with respect to
the lacZ promoter of pBK-CMV. An amplified library was
prepared from approximately 3 million primary clones in
the lambda ZAP Express vector. In vivo excision was then

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2004, 04:30:48 ; Search time 4588 Seconds
(without alignments)
10459.240 Million cell updates/sec

Title: US-09-643-755B-1
Perfect score: 1173
Sequence: 1 atgaactccctaagctctt.....ggtcagctcaagcaatctga 1173

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 20454813386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_mus:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1173	100.0	1173	6	AX088019
2	1173	100.0	3957	6	AX088021
3	56	4.8	1269	4	BOVCHYMOC
4	49	4.2	858	4	AF421165
5	49	4.2	957	4	AF421164
6	49	4.2	1095	4	AF421161
7	49	4.2	1098	6	E00079
8	49	4.2	1098	6	E02341
9	49	4.2	1101	4	AF177290
10	49	4.2	1115	12	SYNPROCA
11	49	4.2	1175	6	AR002347
12	49	4.2	1240	6	AR002347
13	49	4.2	1275	4	BOVCHYMOA
14	49	4.2	1289	6	E00144
15	49	4.2	1291	6	E15836
16	49	4.2	1305	4	BOVCHYMOB
17	49	4.2	1460	6	E00042
18	49	4.2	1460	6	E00295
19	49	4.2	2726	6	E108097
20	49	4.2	2733	6	AR073077
21	49	4.2	2982	6	108098
22	49	3.8	1094	4	BP19786
23	38	3.2	843	4	AF421166
24	38	3.2	1080	4	AF421162
25	38	3.2	1290	6	E00075
26	38	3.2	1291	6	A15633
27	38	3.2	1311	6	E00108
28	37	3.2	119	4	BOVCHYMO8
29	35	3.0	171	4	BOVCHYMO2
30	35	3.0	637	6	B05472
31	35	3.0	744	4	AF421167
32	35	3.0	981	4	AF421163
33	34	2.9	1096	6	BD096860
34	34	2.8	179	6	104059
35	33	2.8	819	6	BD096861
36	33	2.8	1172	4	CDR131677
37	32	2.7	335	4	BOVCHYMO9
38	32	2.7	1292	4	QAPCHY
39	30	2.6	134	4	BOVCHYMO6
40	30	2.6	165	4	BOVCHYMO7
41	29	2.5	1460	6	E00042
42	29	2.5	1460	6	E00295
43	27	2.3	139	4	BOVCHYMO4
44	25	2.1	138	4	BOVCHYMO3
45	23	2.0	220	4	BOVCHYMO5

ALIGNMENTS

RESULT 1
LOCUS AX088019 1173 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 1 from Patent WO0114571.
ACCESSION AX088019
VERSION AX088019.1 GI:13396947
KEYWORDS
SOURCE
ORGANISM
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS
TITLE
1 van Rooijen, G., Keon, R.G., Boothe, J. and Shen, Y.
Commercial production of chymosin in plants

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0114571-A 1 01-MAR-2001;
Sembiosys Genetics Inc. (CA)
FEATURES
source
1. .1173
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
1. .1173
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC34680.1"
/db_xref="GI:13396948"
/translation="MNFLKSPFPYAFLCQYFVAVTHAAETRIPLYYKSLRKALK
EHLLEDFLOKQOYGISSKYSGGEVASVPLTNVLDISOYFGKIYLGTPPEFVLDY
GSSDFWVPSIYCKSNACKNKHORFDPKRSFTQNLGKPLSIHYGTSGMOGILGDTV
SNIVDIQOVTGLSTOEPGVFTYAEFDGLIEMAYPSLASEYSIPVDNMNRLVAD
LFSVYDNRNGOESMLTGAIDPSYTGSLHWPVTYQOYQWFTVDSVTISGVVACEG
GCCAIIIDTSTSLKLVGRSSDILNIQQAIGATONQYGEFDIDCDNLSTMPVTFEINRKM
YPLTPSAVYSQDGFCSTGFSQSENHOKWILGDVFIREFYVSFDRANNLVGLAKAI"

BASE COUNT 299 a 308 c 262 g 304 t
ORIGIN

Query Match 100.0%; Score 1173; DB 6; Length 1173;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACCTCTTAACTCTTCCCTTTCTACGCTTCTCTTGTTCGATCAATCTCGTT 60
DB 1 ATGAACCTCTTAACTCTTCCCTTTCTACGCTTCTCTTGTTCGATCAATCTCGTT 60
QY 61 GCTGTACTACGCTGTGATCACCAGATTCCTCTCAAAAGGTAAGTCTCCCT 120
DB 61 GCTGTACTACGCTGTGATCACCAGATTCCTCTCAAAAGGTAAGTCTCCCT 120
QY 121 AAGCGCTGAAGAACATGACTCTTAGAAGCTTCTTGAGAAACAACAGATGCGATC 180
DB 121 AAGCGCTGAAGAACATGACTCTTAGAAGCTTCTTGAGAAACAACAGATGCGATC 180
QY 181 AGCAGCAGTACTCCGGCTTGGTGAAGTTGCTAGCGTGCCTTACCACTACCTTGAT 240
DB 181 AGCAGCAGTACTCCGGCTTGGTGAAGTTGCTAGCGTGCCTTACCACTACCTTGAT 240
QY 241 AGTCATCTCTTGGGAAGATCTACCTCGGAACCCCGCTCAAGAGTCAACCTCTCTT 300
DB 241 AGTCATCTCTTGGGAAGATCTACCTCGGAACCCCGCTCAAGAGTCAACCTCTCTT 300
QY 301 GATATCTGTCTCTCTGACTTCTGGGTTCCCTCTATCTACTGCAAGCAATGCTGCAAG 360
DB 301 GATATCTGTCTCTCTGACTTCTGGGTTCCCTCTATCTACTGCAAGCAATGCTGCAAG 360
QY 361 AACCAACCAAGATTCGATCCGAGAAAGTCTCACTTCAAGAACTTAGGCAAAACCTTTG 420
DB 361 AACCAACCAAGATTCGATCCGAGAAAGTCTCACTTCAAGAACTTAGGCAAAACCTTTG 420
QY 421 TCTATACATACGATGATAGGATGATGCAAGAACTTAGGATGATACCTGCACTGTC 480
DB 421 TCTATACATACGATGATAGGATGATGCAAGAACTTAGGATGATACCTGCACTGTC 480
QY 481 TCCAACTTGGACATTTCAACAGACAGTGAAGCTTAGCAACCAAGCAAGGATGATGTC 540
DB 481 TCCAACTTGGACATTTCAACAGACAGTGAAGCTTAGCAACCAAGCAAGGATGATGTC 540
QY 541 TTCACTATGACAGAAATTCGATGCGATCTTGGTATGCGATACCCATGCTCGGTCAGAG 600
DB 541 TTCACTATGACAGAAATTCGATGCGATCTTGGTATGCGATACCCATGCTCGGTCAGAG 600
QY 601 TACTGATACCTGTGTTGAACAATGATGAAACGACCACTAGTACCTCAAGACTTGTTC 660
DB 601 TACTGATACCTGTGTTGAACAATGATGAAACGACCACTAGTACCTCAAGACTTGTTC 660
QY 661 TCGGTTTACATGACAGAAATTCGATGCGATCTTGGTATGCGATACCCATGCTCGGTCAGAG 720
DB 661 TCGGTTTACATGACAGAAATTCGATGCGATCTTGGTATGCGATACCCATGCTCGGTCAGAG 720

QY 721 TCTTACTACACAGAAATCTTCACTGCGGTTCCAGTCACTGTGACAGTACTGCGAATTC 780
DB 721 TCTTACTACACAGAAATCTTCACTGCGGTTCCAGTCACTGTGACAGTACTGCGAATTC 780
QY 781 ACTGTGACAGTGTCACTACAGCGGTTGTTGTTGCAATGTAAGTGAATGTCAGCT 840
DB 781 ACTGTGACAGTGTCACTACAGCGGTTGTTGTTGCAATGTAAGTGAATGTCAGCT 840
QY 841 ATCTTGATACCGGATACGTCACAGCTGTGTCAGCCTTAGACGACATTCACATTCAG 900
DB 841 ATCTTGATACCGGATACGTCACAGCTGTGTCAGCCTTAGACGACATTCACATTCAG 900
QY 901 CAAGCTATTGACGCCACACAGAACCGATACGTTGATGACATTAATGCCAACCTT 960
DB 901 CAAGCTATTGACGCCACACAGAACCGATACGTTGATGACATTAATGCCAACCTT 960
QY 961 AGCTACATGCTCTCACTTGTCTTTGATCAACCGCAAGATGATACCACTGACCCCTCC 1020
DB 961 AGCTACATGCTCTCACTTGTCTTTGATCAACCGCAAGATGATACCACTGACCCCTCC 1020
QY 1021 GCCTATACACAGCAGATCAAGGTTCTGCACAGTGAATTCAGATGAGAACCATTC 1080
DB 1021 GCCTATACACAGCAGATCAAGGTTCTGCACAGTGAATTCAGATGAGAACCATTC 1080
QY 1081 CAGAAATGATCTTGGAGATGTCTTCACTTCTGATCTACACGCTCTTTGACAGGCTC 1140
DB 1081 CAGAAATGATCTTGGAGATGTCTTCACTTCTGATCTACACGCTCTTTGACAGGCTC 1140
QY 1141 AACACCTGTTGGGCTAGCTAAAGCAATCTGA 1173
DB 1141 AACACCTGTTGGGCTAGCTAAAGCAATCTGA 1173

RESULT 2
AX088021
LOCUS AX088021 3957 bp DNA linear PAT 17-MAR-2001
DEFINITION Sequence 3 from Patent WO0114571.
ACCESSION AX088021
VERSION AX088021.1 GI:13396949
KEYWORDS
SOURCE
ORGANISM
synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS van Rooijen, G., Keon, R. G., Boothe, J. and Shen, Y.
TITLE Commercial production of chymosin in plants
JOURNAL Patent: WO 0114571-A 3 01-MAR-2001;
Sembiosys Genetics Inc. (CA)
FEATURES
source
1. .3957
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
1554. .2726
/note="Figure 2"
/codon_start=1
/transl_table=11
/protein_id="CAC34681.1"
/db_xref="GI:13396950"
/translation="MNFLKSPFPYAFLCQYFVAVTHAAETRIPLYYKSLRKALK
EHLLEDFLOKQOYGISSKYSGGEVASVPLTNVLDISOYFGKIYLGTPPEFVLDY
GSSDFWVPSIYCKSNACKNKHORFDPKRSFTQNLGKPLSIHYGTSGMOGILGDTV
SNIVDIQOVTGLSTOEPGVFTYAEFDGLIEMAYPSLASEYSIPVDNMNRLVAD
LFSVYDNRNGOESMLTGAIDPSYTGSLHWPVTYQOYQWFTVDSVTISGVVACEG
GCCAIIIDTSTSLKLVGRSSDILNIQQAIGATONQYGEFDIDCDNLSTMPVTFEINRKM
YPLTPSAVYSQDGFCSTGFSQSENHOKWILGDVFIREFYVSFDRANNLVGLAKAI"

BASE COUNT 1263 a 790 c 1295 t
ORIGIN

Query Match 100.0%; Score 1173; DB 6; Length 3957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACCTTCTTAAAGTCTTCCCTTCTTACGCTTCTTGTGTTGGTCAATTAATCTGCT 60
 Db 1554 ATGAACCTTCTTAAAGTCTTCCCTTCTTACGCTTCTTGTGTTGGTCAATTAATCTGCT 1613
 QY 61 GCTGTATCTACAGCTGCTGAGATCAACCCGCAATCTCTCTTCAAAAGTAAGTCTCCGT 120
 Db 1614 GCTGTATCTACAGCTGCTGAGATCAACCCGCAATCTCTCTTCAAAAGTAAGTCTCCGT 1673
 QY 121 AAGGCGCTGAAGGAACATGACCTTCTTGAAGCTTCTTGAAGAAACAAGATGGCATC 180
 Db 1674 AAGGCGCTGAAGGAACATGACCTTCTTGAAGCTTCTTGAAGAAACAAGATGGCATC 1733
 QY 181 AGCAGCAAGTACTCCGCTGCTGAGAGTCTAGCGTCACTTACCACTTACCTTAT 240
 Db 1734 AGCAGCAAGTACTCCGCTGCTGAGAGTCTAGCGTCACTTACCACTTACCTTAT 1793
 QY 241 AGTCATATCTTTGGAGAGATCTACTCGGAAACCCGCTCAAGAGTTCAACGTTCTCTT 300
 Db 1794 AGTCATATCTTTGGAGAGATCTACTCGGAAACCCGCTCAAGAGTTCAACGTTCTCTT 1853
 QY 301 GATACCTGTTCTCTGATTTCTGGTTCCCTCTTATCTTCTGCAAGAAATGCTTGCAG 360
 Db 1854 GATACCTGTTCTCTGATTTCTGGTTCCCTCTTATCTTCTGCAAGAAATGCTTGCAG 1913
 QY 361 AACCAACAAAGTTGATTCGAGAAAGTCTGTCACCTTCCAGAACTTAGGCAAAACCTTG 420
 Db 1914 AACCAACAAAGTTGATTCGAGAAAGTCTGTCACCTTCCAGAACTTAGGCAAAACCTTG 1973
 QY 421 TCTATATACCTACGAGTACAGTAGACATGCAAGAAATCTTAGGTATGATACCTGCTGTC 480
 Db 1974 TCTATATACCTACGAGTACAGTAGACATGCAAGAAATCTTAGGTATGATACCTGCTGTC 2033
 QY 481 TCCAACTTTGAGAGATTTCAACAGACATGAGACTTACGACCCCAAGAACCAAGTATGTC 540
 Db 2034 TCCAACTTTGAGAGATTTCAACAGACATGAGACTTACGACCCCAAGAACCAAGTATGTC 2093
 QY 541 TTCACCTTATGAGAAATGATGAGATCTGCTGATGAGATACCCATGCTGCTGCTGAGAG 600
 Db 2094 TTCACCTTATGAGAAATGATGAGATCTGCTGATGAGATACCCATGCTGCTGCTGAGAG 2153
 QY 601 TACTGATATCTGTTGTTGACAAACATGATGAACCGACACTTAGTACTCAAGACTTGTTC 660
 Db 2154 TACTGATATCTGTTGTTGACAAACATGATGAACCGACACTTAGTACTCAAGACTTGTTC 2213
 QY 661 TCGGTTTACATGAGAGAGATGCGCAGAGAGACATGCTCAGCTTGGAGCTATTGATCA 720
 Db 2214 TCGGTTTACATGAGAGAGATGCGCAGAGAGACATGCTCAGCTTGGAGCTATTGATCA 2273
 QY 721 TCCCTACTACAGAGATCTCTTCACTGGGTTCCAGTCACTGTGACAGCACTTACCTGGCAATTC 780
 Db 2274 TCCCTACTACAGAGATCTCTTCACTGGGTTCCAGTCACTGTGACAGCACTTACCTGGCAATTC 2333
 QY 781 ACTGTGACAGTGTTCACCATCAGCGGTGTGTTGTGATGTAAGTGAATGTCAACT 840
 Db 2334 ACTGTGACAGTGTTCACCATCAGCGGTGTGTTGTGATGTAAGTGAATGTCAACT 2393
 QY 841 ATCTTGTATCCGGTACGTCCAGCTGTGCTGCACTTACAGCGACATTTCTCAACATTGAG 900
 Db 2394 ATCTTGTATCCGGTACGTCCAGCTGTGCTGCACTTACAGCGACATTTCTCAACATTGAG 2453
 QY 901 CAAGCTATTTGAGCCACACAGAACAGTACGTTGATTTGACATGATTTGGAACAACCTT 960
 Db 2454 CAAGCTATTTGAGCCACACAGAACAGTACGTTGATTTGACATGATTTGGAACAACCTT 2513
 QY 961 AGCTACATGCTTACAGTGTCTTTGATGATCAACGCAAGATGTACCACTGACCCCTTCC 1020
 Db 2514 AGCTACATGCTTACAGTGTCTTTGATGATCAACGCAAGATGTACCACTGACCCCTTCC 2573
 QY 1021 GCTTATACAGCGCAGATTCAGGGTTCGACCACTGATTTCAAGTGAAGAACATTTC 1080
 Db 2574 GCTTATACAGCGCAGATTCAGGGTTCGACCACTGATTTCAAGTGAAGAACATTTC 2633

QY 1081 CAGAAATGATCTTTGGAGATGTTGTTCTTCTGATGATCTACAGCGTCTTTGACAGGGCC 1140
 Db 2634 CAGAAATGATCTTTGGAGATGTTGTTCTTCTGATGATCTACAGCGTCTTTGACAGGGCC 2693
 QY 1141 AACCACTCTGTTGGGCTTAGCTTAAGCAATCTGA 1173
 Db 2694 AACCACTCTGTTGGGCTTAGCTTAAGCAATCTGA 2726

RESULT 3

BOVCHYMOC 1269 bp mRNA linear MAM 28-FEB-1994
 LOCUS Bovine chymosin C (rennin) mRNA, complete cds.
 DEFINITION J00004
 VERSION J00004.1 GI:162861
 KEYWORDS chymosin; chymosin C; protease; rennin.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus

REFERENCE 1 (bases 1 to 1269)
 AUTHORS Nishimori, K., Kawaguchi, Y., Hidaka, M., Uozumi, T. and Beppu, T.
 TITLE Nucleotide sequence of calf prorennin cDNA cloned in *Escherichia coli*.
 JOURNAL J. Biochem. 91 (3), 1085-1088 (1982)
 MEDLINE 82189915
 PUBMED 6804449

COMMENT Original source text: Bos taurus calf fourth stomach mucosa cDNA to mRNA.

chymosin is the major proteolytic enzyme in the fourth stomach of the unweaned calf. Two chromatographically different forms, a and b, of the enzyme and its precursor are known and a third form, represented by this sequence, seems likely. The presence of a termination codon (bases 27-29) within the prepropeptide coding sequence implies that this sequence might be a pseudogene mRNA, or there may be an error in the cDNA. The amino acid sequence deduced from this nucleotide sequence differs at three sites from bovine chymosin a (see bovyhymoc) and at four sites from bovine chymosin b (see bovyhymob).

FEATURES

source Location/Qualifiers
 1..1269
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="Fourth stomach mucosa"
 /dev_stage="calf"
 <1..1269
 /product="chymoc mRNA"
 402..1130
 /note="precursor"
 /codon_start=1
 /product="chymosin C"
 /protein_id="AAA30449.1"
 /db_xref="GI:457097"
 /translation="MGGILGVDPTVTSNIVDIOGTGASIQEPDVFVTAEPDILIM
 AYSLASRYSLIPVFDNMNRHLVAQDLFSYMDNCGESMLTGLADIPSYTSGLHW
 PVTVQQTWQTVDSVLTISGVVACGGCQALIDGTSLVGPSSDIINIOQATIGON
 QYDFEDIDCNLSYMPVFEINGKIVLPISAYTGDDGFCCTGFPSEHNSOKMILG
 DVFIREYYSVFDNRANIVGLAKTI"

BASE COUNT 299 a 381 c 321 g 268 t
 ORIGIN

Query Match 4.8%; Score 56; DB 4; Length 1269;
 Best Local Similarity 100.0%; Pred. No. 7.3e-18;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 760 GTGAGAGATGATCGCAATTCATGTGAGACAGTCAACATCAGCGGTGTTGT 815
 Db 717 GTGAGAGATGATCGCAATTCATGTGAGACAGTCAACATCAGCGGTGTTGT 772
 RESULT 4

AF421165 858 bp mRNA linear MAM 01-APR-2002
LOCUS Bos taurus chymosin precursor, mRNA, partial cds, alternatively
spliced.
DEFINITION
ACCESSION AF421165
VERSION AF421165.1 GI:19851897
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS Zinovieva,N., McCoy,B., Mueller,M. and Brem,G.
TITLE Multiple splicing forms of bovine prochymosin mRNA, their
identification and characterization
JOURNAL
AUTHORS Zinovieva,N., McCoy,B., Mueller,M. and Brem,G.
TITLE Direct Submission
JOURNAL Submitted (19-SEP-2001) Institute of Animal Breeding and Genetics,
WU, Veterinaerplatz 1, Vienna 1210, Austria
FEATURES
source
1..858
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="fourth stomach mucosa"
/dev_stage="calf"
/EC_number="3.4.23.4"
/note="rennin; preproprotein; alternatively spliced; lacks
exons 3, 4 and 8"
/codon_start=2
/product="chymosin precursor"
/protein_id="AAL99910.1"
/db_xref="GI:19851898"
/translation="VLLAVFALSGTEIRIPLYGKSLRKALKEHGLLEDFLOKQY
GISKTSIGREYASVPLNTYLDVSNIVDIQTGTSLQEPGDVFTYAEFGILGMAYP
SLASEKSIPIVFDNMNRHLVAQDLPSVYMDRNGQESMLTLGALDPSYTGSLHWVPT
VOOYQFIVDSYTIISGVVACGCGCALDITGTSKLVGPSDIINLQALIGATONQYG
EDQFCTSGFQSENHSOKWILGDFIREYYSVFPDRAANLVGLAKAI"
BASE COUNT 193 a 255 c 241 g 169 t
ORIGIN
Query Match 4.2%; Score 49; DB 4; Length 858;
Best Local Similarity 100.0%; Pred. No. 3.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 656 TGTTCGCGTTTACATGACAGGAATGCCAGAGACATGCTCAGCCT 704
Db 378 TGTTCGCGTTTACATGACAGGAATGCCAGAGACATGCTCAGCCT 426
RESULT 5
AF421164 957 bp mRNA linear MAM 01-APR-2002
LOCUS Bos taurus chymosin precursor, mRNA, partial cds, alternatively
spliced.
DEFINITION
ACCESSION AF421164
VERSION AF421164.1 GI:19851895
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS Zinovieva,N., McCoy,B., Mueller,M. and Brem,G.
TITLE Multiple splicing forms of bovine prochymosin mRNA, their
identification and characterization
JOURNAL
AUTHORS Zinovieva,N., McCoy,B., Mueller,M. and Brem,G.
TITLE Direct Submission
JOURNAL Submitted (19-SEP-2001) Institute of Animal Breeding and Genetics,
WU, Veterinaerplatz 1, Vienna 1210, Austria
FEATURES
source
1..1095
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="fourth stomach mucosa"
/dev_stage="calf"
/EC_number="3.4.23.4"
/note="rennin; preproprotein; alternatively spliced; lacks
exon 8"
/product="chymosin precursor"

AUTHORS Zinovieva,N., McCoy,B., Mueller,M. and Brem,G.
TITLE Direct Submission
JOURNAL Submitted (19-SEP-2001) Institute of Animal Breeding and Genetics,
WU, Veterinaerplatz 1, Vienna 1210, Austria
FEATURES
source
1..957
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="fourth stomach mucosa"
/dev_stage="calf"
/EC_number="3.4.23.4"
/note="rennin; preproprotein; alternatively spliced; lacks
exons 3 and 4"
/codon_start=2
/product="chymosin precursor"
/protein_id="AAL99909.1"
/db_xref="GI:19851896"
/translation="VLLAVFALSGTEIRIPLYGKSLRKALKEHGLLEDFLOKQY
GISKTSIGREYASVPLNTYLDVSNIVDIQTGTSLQEPGDVFTYAEFGILGMAYP
SLASEKSIPIVFDNMNRHLVAQDLPSVYMDRNGQESMLTLGALDPSYTGSLHWVPT
VOOYQFIVDSYTIISGVVACGCGCALDITGTSKLVGPSDIINLQALIGATONQYG
EDIDCDNLSTYMPVTFEINGKMPVLTPEATYSDOPGFCSTSGFQSENHSOKWILGDF
IREYYSVFPDRAANLVGLAKAI"
BASE COUNT 218 a 288 c 260 g 191 t
ORIGIN
Query Match 4.2%; Score 49; DB 4; Length 957;
Best Local Similarity 100.0%; Pred. No. 3.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 656 TGTTCGCGTTTACATGACAGGAATGCCAGAGACATGCTCAGCCT 704
Db 378 TGTTCGCGTTTACATGACAGGAATGCCAGAGACATGCTCAGCCT 426
RESULT 6
AF421161 1095 bp mRNA linear MAM 01-APR-2002
LOCUS Bos taurus chymosin precursor, mRNA, partial cds, alternatively
spliced.
DEFINITION
ACCESSION AF421161
VERSION AF421161.1 GI:19851889
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS Zinovieva,N., McCoy,B., Mueller,M. and Brem,G.
TITLE Multiple splicing forms of bovine prochymosin mRNA, their
identification and characterization
JOURNAL
AUTHORS Zinovieva,N., McCoy,B., Mueller,M. and Brem,G.
TITLE Direct Submission
JOURNAL Submitted (19-SEP-2001) Institute of Animal Breeding and Genetics,
WU, Veterinaerplatz 1, Vienna 1210, Austria
FEATURES
source
1..1095
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="fourth stomach mucosa"
/dev_stage="calf"
/EC_number="3.4.23.4"
/note="rennin; preproprotein; alternatively spliced; lacks
exon 8"
/product="chymosin precursor"


```

/protein_id="AAL99906.1"
/db_xref="GI:19851890"
/translation="VLAIFALISQCTEITRILPYLKGSLRKALKEHGLEDEPLQKQY
GISKXSGFGEVASVPLTNVLDIOYFKIYGTPOEFTVLEDSDFWWSYICKS
NGCKKHORFDPKRSSTPONIGKPLSLIHGTGSMOGLIGVDVTVSNIVDIOCTVGLST
QERGVFTYAEFDGILGMAYPSLASEYSLPFEDNMNHNHVAQDIFSYIMDNQESM
LTLGALIDSYYTSLHWVPTVQYQWFTVDSVTLISGVVACEGCAQLIDTSTKL
GPSSDILNIQQALIGATQNOQGEBOGCTCSQSEHNSQKMLIGDVFIREYYSVFDNRAN
"NLVGLAKK1"

BASE COUNT      246 a      333 c      298 g      218 t
ORIGIN
Query Match      4.2%; Score 49; DB 4; Length 1095;
Best Local Similarity 100.0%; Pred. No. 3.8e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 656 TGTTCGCGTTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCT 704
      |||||||
Db 615 TGTTCGCGTTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCT 663

RESULT 7
LOCUS      E00079      1098 bp      RNA      linear      PAT 29-SEP-1997
DEFINITION CDNA sequence encoding calf prorennin.
ACCESSION  E00079
VERSION    E00079.1 GI:2168383
KEYWORDS   JP 1983032896-A/1.
SOURCE     unidentified
ORGANISM   unidentified
REFERENCE   1 (bases 1 to 1098)
AUTHORS    Beppu, T., Uozumi, T. and Nishimori, K.
TITLE      CONJUGATED PLASMOD AND MICROORGANISM CONTAINING THE SAME
JOURNAL    Patent: JP 1983032896-A 1 25-FEB-1983;
           BEPPU TERUHIKO
COMMENT     OS      calf
           PN      JP 1983032896-A/1
           PD      25-FEB-1983
           PF      24-AUG-1981 JP 1981131631
           PI      BEPPU TERUHIKO, UOZUMI TAKESHI, NISHIMORI KATSUHIKO PC
           C07H21/04, C12N1/00, C12N15/00, C12N1/00, C12R1/19; CC
           strandedness: Double;
           CC      topology: linear;
           CC      hypothetical: No;
           CC      anti-sense: No;
           CC      *source: tissue=mucoous membrane of forth stomach; CC
           FH      *source: clone=PCR 100 1;
           FH      Key      Location/Qualifiers
           FT      CDS      1..1098
                        /product="calf prorennin",
                        /location/Qualifiers
                        1..1098
                        /organism="unidentified"
                        /mol_type="genomic RNA"
                        /db_xref="taxon:32644"
BASE COUNT      252 a      327 c      300 g      219 t
ORIGIN
Query Match      4.2%; Score 49; DB 6; Length 1098;
Best Local Similarity 100.0%; Pred. No. 3.8e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 656 TGTTCGCGTTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCT 704
      |||||||
Db 581 TGTTCGCGTTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCT 629

RESULT 8
LOCUS      E02341      1098 bp      RNA      linear      PAT 29-SEP-1997
DEFINITION CDNA sequence coding for infant bovine prorennin.

```

```

ACCESSION  E02341
VERSION    E02341.1 GI:2170576
KEYWORDS   JP 1990109984-A/1.
SOURCE     Bos taurus (cow)
ORGANISM   Bos taurus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
           Bovidae; Bovinae; Bos.
REFERENCE   1 (bases 1 to 1098)
AUTHORS    Beppu, T., Uozumi, T. and Nishimori, K.
TITLE      COMPLEX PLASMOD AND MICROORGANISM CONTAINING THE SAME
JOURNAL    Patent: JP 1990109984-A 1 23-APR-1990;
           BEPPU TERUHIKO
COMMENT     OS      Bovine
           PN      JP 1990109984-A/1
           PD      23-APR-1990
           PF      01-DEC-1988 JP 1988302176
           PI      BEPPU TERUHIKO, UOZUMI TAKESHI, NISHIMORI KATSUHIKO PC
           C12N15/59, C12N1/21;
           CC      strandedness: Single;
           CC      topology: linear;
           CC      hypothetical: No;
           CC      anti-sense: No;
           CC      *source: tissue=Stomach mucosa;
           FH      Key      Location/Qualifiers
           FT      CDS      1..1098
                        /product="Infant bovine prorennin",
                        /location/Qualifiers
                        1..1098
                        /organism="Bos taurus"
                        /mol_type="genomic RNA"
                        /db_xref="taxon:9913"
BASE COUNT      252 a      327 c      300 g      219 t
ORIGIN
Query Match      4.2%; Score 49; DB 6; Length 1098;
Best Local Similarity 100.0%; Pred. No. 3.8e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 656 TGTTCGCGTTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCT 704
      |||||||
Db 581 TGTTCGCGTTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCT 629

RESULT 9
LOCUS      AF177290      1101 bp      mRNA      linear      MAM 24-JAN-2000
DEFINITION Bubalus arnee bubalis prothymosin mRNA, complete cds.
ACCESSION  AF177290
VERSION    AF177290.1 GI:6739579
KEYWORDS
SOURCE     SOURCE
           Bubalus bubalis (water buffalo)
           Bubalus bubalis
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
           Bovidae; Bovinae; Bubalus.
           1 (bases 1 to 1101)
           Batish, V.K., Mukhopadhyay, U.K., Mohanty, A.K., Grover, S. and
           Kuipers, O.P.
           Direct Submission
           Submitted (12-AUG-1999) Molecular Biology Unit, Animal
           Biotechnology Centre, National Dairy Research Institute, G.T.Road,
           Karnal, Haryana 132001, India
           Location/Qualifiers
           1..1101
           /organism="Bubalus bubalis"
           /mol_type="mRNA"
           /sub_species="bubalis"
           /db_xref="taxon:89462"
           1..1101
           /codon_start=1
           /product="prothymosin"

```

/protein_id="AAF27315.1"
/db_xref="GI:6739580"
/translation="MAETIRIPLCCKSKIRKALKENGLEDFLOKQYGVSSKYSFG
EVAASVPLTNVLDQYFGKILGTGPQEFVLPEFTSSDFWVBSIYCKSNACKNHQRFD
PKRSYFONIGKRLSIRYGTSMQGLIGDTATVSNIVDIQVGLSTQEPGVPTYA
EPDGLGMAYPSILASEYSIPVFNMMNRHLVADLPSVMDRGSRMLTLGADPYSY
YVGSILHWVAVTQYQWFTVDSITISGVAVACGGCOALIDGTSTLGVSPSILNIO
QALGATONOGYGFEDIDCNLSVMPYVSLNGLMYLTSAIYSQDQGCCTSGFQSEN
RSCQMLGCVFIHFHYISVFDNRANLVGLAKAT"

BASE COUNT 244 a 336 c 306 g 215 t

Query Match 4.2%; Score 49; DB 4; Length 1101;
Best Local Similarity 100.0%; Pred. No. 3.8e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 656 TGTTCCTCGTTTACATGACAGAAATGCCAGAGAGCATGCTCACGCT 704
Db 584 TGTTCCTCGTTTACATGACAGAAATGCCAGAGAGCATGCTCACGCT 632

RESULT 10
LOCUS SYNPROCAA 1115 bp DNA linear SYN 27-APR-1993
DEFINITION Synthetic bovine prochymosin A protein region.
ACCESSION M22593 M18758
VERSION M22593.1 GI:209139
KEYWORDS prochymosin.
SOURCE synthetic construct
ORGANISM synthetic construct
COMMENT 1 (bases 1 to 1115)
AUTHORS Wosnick,M.A., Barnett,R.W., Vicentini,A.M., Erfle,H., Elliott,R.,
Summer-Smith,M., Mantel,N. and Davies,R.W.
TITLE Rapid construction of large synthetic genes: total chemical
synthesis of two different versions of the bovine prochymosin gene
JOURNAL Gene 60 (1), 115-127 (1987)
MEDLINE 88152494
PUBMED 3126097
COMMENT Original source text: Synthetic DNA.
FEATURES
source location/Qualifiers
1..1115
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

BASE COUNT 254 a 332 c 302 g 227 t

ORIGIN 1 bp downstream of HindIII site.

Query Match 4.2%; Score 49; DB 12; Length 1115;
Best Local Similarity 100.0%; Pred. No. 3.8e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 656 TGTTCCTCGTTTACATGACAGAAATGCCAGAGAGCATGCTCACGCT 704
Db 580 TGTTCCTCGTTTACATGACAGAAATGCCAGAGAGCATGCTCACGCT 628

RESULT 11
LOCUS 104058 1175 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 5 from Patent EP 0123928.
ACCESSION 104058
VERSION 104058.1 GI:591912
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
COMMENT Unclassified.
1 (bases 1 to 1175)
AUTHORS Cashion,L.M., McCaman,M.T., Rice,C.W. and Sias,S.R.
TITLE Recombinant DNA coding for a polypeptide displaying milk clotting
activity
JOURNAL Patent: EP 0123928-A2 5 07-NOV-1984;
FEATURES Location/Qualifiers

source 1..1175
/organism="unknown"

BASE COUNT 263 a 356 c 318 g 238 t

ORIGIN

Query Match 4.2%; Score 49; DB 6; Length 1175;
Best Local Similarity 100.0%; Pred. No. 3.8e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 656 TGTTCCTCGTTTACATGACAGAAATGCCAGAGAGCATGCTCACGCT 704
Db 649 TGTTCCTCGTTTACATGACAGAAATGCCAGAGAGCATGCTCACGCT 697

RESULT 12
LOCUS AR002347 1240 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 2 from patent US 5741665.
ACCESSION AR002347
VERSION AR002347.1 GI:3963901
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
COMMENT Unclassified.
1 (bases 1 to 1240)
AUTHORS Kato,E.K. and Stuart,W.Dorsey.
TITLE Light-regulated promoters for production of heterologous proteins
in filamentous fungi
JOURNAL Patent: US 5741665-A 2 21-APR-1998;
FEATURES Location/Qualifiers
source 1..1240
/organism="unknown"

BASE COUNT 274 a 374 c 339 g 253 t

ORIGIN

Query Match 4.2%; Score 49; DB 6; Length 1240;
Best Local Similarity 100.0%; Pred. No. 3.7e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 656 TGTTCCTCGTTTACATGACAGAAATGCCAGAGAGCATGCTCACGCT 704
Db 700 TGTTCCTCGTTTACATGACAGAAATGCCAGAGAGCATGCTCACGCT 748

RESULT 13
LOCUS BOVCHYM0A 1275 bp mRNA linear MAM 19-DEC-2002
DEFINITION Bos taurus preprochymosin a mRNA, complete cds.
ACCESSION J00002
VERSION J00002.1 GI:162857
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 1275)
AUTHORS Molir,D., Mao,J., Schumm,J.W., Vovis,G.F., Alford,B.L. and
Taunton-Rigby,A.
TITLE Molecular cloning and characterization of double-stranded cDNA
coding for bovine chymosin
JOURNAL Gene 19 (1), 127-138 (1982)
MEDLINE 83054629
PUBMED 6183168
COMMENT chymosin (rennin) is the major proteolytic enzyme in the fourth
stomach of the unweaned calf. two chromatographically distinct
forms are known and a third seems likely (see bovcnymob and
bovcnymoc). this sequence has been tentatively identified as
chymosin a; it differs from chymosin b by only two amino acids and
from chymosin c by another amino acid. [1] argues that the
different chymosins are probably polymorphic variants of a single
chymosin gene.
FEATURES Location/Qualifiers

source 1..1275
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="fourth stomach mucosa"
21..1166
/note="remnin"
/codon_start=1
/product="preprochymosin a"
/protein_id="AAA30447.1"
/db_xref="GI:162858"
/translation="MRCVVLAVPALVPSGAEITRIPIYKSKSLKALKEHLEDEL
QKQYGISKYSGEFVASVPLTNLYDSQYFGLTYLGPQEFVLFDPTGSSDFWVPS
LYCKSNACKNQRDPKRSSTFONLGRPLSIHYGSMQGLIYDPTVTSNILDIOQT
VGLSTQEPDVFTYAEFDGILGMAYPSLASEYSIPVPMNMRHLVAODLFSVMDRN
GOESMLTGLAIDPCVYGLSHWPTVVOYQVPTVDSVTISGVAVAGCGCOALIDTG
TSKLVGPSDLINIOAIGATQNGYDERDIDCDNISYPTVFEFLNGKMYPLTTSAYT
SODQFCTSGFQSESHSQWILGDVFINETYSVFDNRANNVGLAKAT"
24..68
/note="presequence"
/product="chymosin a"
195..1163
/product="chymosin a"
293 a 391 c 336 g 255 t

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 4.2%; Score 49; DB 4; Length 1275;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 656 TGTTCGCGTTTACATGACAGAAATGGCCAGAGAGCATGCTCAGCT 704
|||||
Db 649 TGTTCGCGTTTACATGACAGAAATGGCCAGAGAGCATGCTCAGCT 697
|||||

RESULT 14
E00144 1289 bp RNA linear PAT 29-SEP-1997
LOCUS CDNA encoding bovine calf chymosin.
DEFINITION E00144
ACCESSION E00144
VERSION E00144.1 GI:2168443
KEYWORDS JP 1984021392-A/1.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1289)
AUTHORS Chiyaarusu, E.B.
TITLE CATTLE KIMOCIN
JOURNAL Patent: JP 1984021392-A 1 03-FEB-1984;
GENEX CORP
OS bovine calf
PN JP 1984021392-A/1
PD 03-FEB-1984
PF 30-JUN-1983 JP 1983119481
PR 01-JUL-1982 US 82 394433, 13-APR-1983 US 83 484539 PI
CHITAAARUSU EI BASURETSUTO
PC C12N15/00, C07H21/04, C12N1/20, C12P19/34, C12P21/02, (C12N15/00,
PC C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: tissue_type=stomach mucosa;
FH key
FH Location/Qualifiers
FT mat_peptide 169..1113
FT /product='bovine calf chymosin' FT CDS
FT <1..1116
FT /product='precursor polypeptide'.

FEATURES
source
1..1289
/organism="unclassified"
/mol_type="genomic RNA"
/db_xref="taxon:32644"

BASE COUNT 311 a 388 c 329 g 260 t 1 others
ORIGIN

Query Match
Best Local Similarity 4.2%; Score 49; DB 6; Length 1289;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 656 TGTTCGCGTTTACATGACAGAAATGGCCAGAGAGCATGCTCAGCT 704
|||||
Db 623 TGTTCGCGTTTACATGACAGAAATGGCCAGAGAGCATGCTCAGCT 671
|||||

RESULT 15
A15836 1291 bp DNA linear PAT 16-FEB-1994
LOCUS chymosin gene.
DEFINITION A15836
ACCESSION A15836
VERSION A15836.1 GI:488959
KEYWORDS Bos taurus (cow)
SOURCE Bos taurus
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 1291)
AUTHORS Simons, A.F.M. and De Vos, W.M.
TITLE DNA fragments, containing a lactic acid bacterium-specific
regulator region for the expression of genes coding for normally
heterologous proteins
Patent: EP 0307011-A 5 15-MAR-1989;
NEDERLANDS INSTITUUT VOOR ZUIVELANDERBOEK
Location/Qualifiers
source
1..1291
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"

BASE COUNT 323 a 382 c 328 g 257 t 1 others
ORIGIN

Query Match
Best Local Similarity 4.2%; Score 49; DB 6; Length 1291;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 656 TGTTCGCGTTTACATGACAGAAATGGCCAGAGAGCATGCTCAGCT 704
|||||
Db 609 TGTTCGCGTTTACATGACAGAAATGGCCAGAGAGCATGCTCAGCT 657
|||||

Search completed: February 5, 2004, 07:24:19
Job time : 4594 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2004, 03:13:23 ; Search time 390 Seconds
(without alignments)
8119.079 Million cell updates/sec

Title: US-09-643-755B-1

Perfect score: 1173
Sequence: 1 atgaactctctaagttctt.....gctagctaaagcaatctga 1173

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1173	100.0	1173	22	AA500569
2	1173	100.0	3957	22	AA500570
3	49	4.2	1098	4	AA30063
4	49	4.2	1098	11	AA04683
5	49	4.2	1175	5	AA40285
6	49	4.2	1175	13	AA020949
7	49	4.2	1210	12	AA014051
8	49	4.2	1240	16	AA03006

9	49	4.2	1278	5	AA40055
10	49	4.2	1289	4	AA30022
11	49	4.2	1291	10	AA91157
12	49	4.2	1314	4	AA30049
13	49	4.2	1460	3	AA20043
14	49	4.2	1460	5	AA40180
15	49	4.2	2727	10	AA91188
16	49	4.2	2733	20	AA206463
17	49	4.2	2733	24	AB53073
18	49	4.2	2982	10	AA91185
19	38	3.2	1290	4	AA30209
20	35	3.0	637	14	AA049459
21	34	2.9	1096	20	AA63966
22	33	2.8	179	5	AA40296
23	33	2.8	186	13	AA020950
24	33	2.8	819	20	AA63967
25	32	2.7	107	12	AA014777
26	29	2.5	1460	3	AA20043
27	29	2.5	1460	5	AA40180
28	23	2.0	609	24	AB06309
29	23	2.0	1083	21	AA044658
30	23	2.0	1140	24	AA597159
31	23	2.0	1238	24	AB58369
32	23	2.0	1521	24	AB214440
33	23	2.0	1814	21	AA037875
34	20	1.7	828	11	AA003224
35	20	1.7	846	11	AA003223
36	20	1.7	1104	9	AA080001
37	20	1.7	1143	5	AA40214
38	20	1.7	1146	24	AB53735
39	20	1.7	116624	19	AAV52850
40	19	1.6	231	23	AA57374
41	19	1.6	556	24	ABV67911
42	19	1.6	689	19	AAV59757
43	19	1.6	689	24	AB573750
44	19	1.6	1353	21	AA239862
45	19	1.6	1751	19	AAV59610

ALIGNMENTS

RESULT 1	AA500569	standard; DNA; 1173 BP.
ID	AA500569	
XX	AA500569;	
AC	14-MAY-2001	(first entry)
XX		
DT		
XX		
DE		Bovine pre-pro-chymosin DNA sequence.
XX		
KW		Chymosin; transcription regulator; terminator sequence; soybean; corn;
KW		pre-pro-chymosin; rape seed; sunflower; cotton; tobacco; alfalfa; wheat;
KW		barley; oats; soyghum; Arabidopsis thaliana; potato; flax; linseed; rice;
KW		safflower; oil palm; ground nut; Brazil nut; coconut; castor; coriander;
KW		squash; jojoba; ds.
OS		Bos sp.
XX		
XX		
Key		Location/Qualifiers
FT		1..1173
FT		/tag= a
FT		/product= "Bovine chymosin"
FT		sig_peptide
FT		1..78
FT		/tag= b
FT		misc_feature
FT		79..201
FT		/tag= c
FT		/note= "Pro sequence"
FT		mat_peptide
FT		202..1170
FT		/tag= d
FT		/product= "Mature bovine chymosin"
XX		

Sequence of prochy
Sequence of veal c
Cloned sequence of
cDNA sequence corr
Pre-prorennin-A ge
Sequence of recomb
BamHI/SalI insert
2.7 Kbp HindIII fr
DNA encoding oleos
BamHI insert from
Sequence of prepro
Prochymosin gene 5
DNA encoding the f
Sequence coding fo
DNA coding for a r
DNA encoding His-c
Renin - casein co
Pre-prorennin-A ge
Sequence of recomb
Arabidopsis thalia
Arabidopsis thalia
Human aspartyl pro
Protein modificati
Arabidopsis thalia
Arabidopsis thalia
DNA fragment of PA
DNA fragment of PA
Optimised prochymo
Gene encoding (pre
Aspergillus gene e
Human eyal gene co
cDNA #50 encoding
Human secreted pro
Human cDNA #2 for
Plasmodium ovale p
Human secreted pro

```

PN MO200114571-A1.
XX
PD 01-MAR-2001.
XX
PF 23-AUG-2000; 2000MO-CA00975.
XX
PR 23-AUG-1999; 99US-0378696.
XX
PA (SEMB-) SEMBIOSIS GENETICS INC.
XX
PI Van Rooijen G, Keon RG, Boothe J, Shen Y;
XX
DR P-PSDB; AAU00536.
XX
PT Producing chymosin in seeds of plants such as rice, flax, rape seed, by
XX transforming plant cell with a nucleic acid encoding chymosin operably
XX linked to transcription regulator and terminator sequences -
XX
PS Claim 9; Fig 1; 56pp; English.
XX
CC The sequence represents a DNA which encodes a bovine chymosin
CC polypeptide. Chymosin can be produced in a plant seed through
CC introduction of a chimeric nucleic acid molecule, comprising a nucleic
CC acid sequence encoding a chymosin polypeptide operatively linked to
CC transcription regulator and terminator sequences, into a plant cell. The
CC sequences are useful for producing plant seeds, in particular seeds of
CC soybean, rape seed, sunflower, cotton, corn, tobacco, alfalfa, wheat,
CC barley, oats, sorghum, Arabidopsis thaliana, potato, flax/linseed,
CC safflower, oil palm, groundnut, Brazil nut, coconut, castor, coriander,
CC squash, jojoba and rice.
XX
SQ Sequence 1173 BP; 299 A; 308 C; 262 G; 304 T; 0 other;

Query Match      100.0%; Score 1173; DB 22; Length 1173;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACCTCTTAACTCTTCCCTTCTAGAGCTTCTCTTGGTCAATACCTCGTT 60
DB 1 ATGAACCTCTTAACTCTTCCCTTCTAGAGCTTCTCTTGGTCAATACCTCGTT 60
QY 61 GCTGTTACTCAGCGTGTGAGATCAGCCGATTCCTCTCAAAAGTAAGTCTCTCGT 120
DB 61 GCTGTTACTCAGCGTGTGAGATCAGCCGATTCCTCTCAAAAGTAAGTCTCTCGT 120
QY 121 AAGGCGCTGAAGAACATGAGACTTCTAGAGACTTCTTGAGAAACAACAGTAGCATC 180
DB 121 AAGGCGCTGAAGAACATGAGACTTCTAGAGACTTCTTGAGAAACAACAGTAGCATC 180
QY 181 AGCAGCAGTAGTCTCCGCTTGTGAGAGTGTGAGCGTGCCTTACCACTACCTTGAT 240
DB 181 AGCAGCAGTAGTCTCCGCTTGTGAGAGTGTGAGCGTGCCTTACCACTACCTTGAT 240
QY 241 AGTCAATCTTGGAGAGATCTACTGTGAAACCCGCTCAAGAGTTCAACCTTCTTT 300
DB 241 AGTCAATCTTGGAGAGATCTACTGTGAAACCCGCTCAAGAGTTCAACCTTCTTT 300
QY 301 GATACCTGTTCTCTGACTTCTGGGTTCCCTCTACTACTGCAAGCAATGCTTGCAAG 360
DB 301 GATACCTGTTCTCTGACTTCTGGGTTCCCTCTACTACTGCAAGCAATGCTTGCAAG 360
QY 361 AACCACCAAAAGATTGATCCGAGAAAGTGTGCACTTCCAGAACTTAGGCAAACTTTG 420
DB 361 AACCACCAAAAGATTGATCCGAGAAAGTGTGCACTTCCAGAACTTAGGCAAACTTTG 420
QY 421 TCTATACACTACCGTATAGAGTAGATGCAAGAGATCTTAGGCTATGATCCGTACCTGC 480
DB 421 TCTATACACTACCGTATAGAGTAGATGCAAGAGATCTTAGGCTATGATCCGTACCTGC 480
QY 481 TCCAACTTGTGACATTTCAACAGACAGTAGAGCTTGAACCCCAAGAACAGGTGATGTC 540
DB 481 TCCAACTTGTGACATTTCAACAGACAGTAGAGCTTGAACCCCAAGAACAGGTGATGTC 540

```

```

QY 541 TTCACCTATGCAAAATTGAGTGGCATCTTGGTATGAGCATACCCATCGCTCGGTCAAG 600
DB 541 TTCACCTATGCAAAATTGAGTGGCATCTTGGTATGAGCATACCCATCGCTCGGTCAAG 600
QY 601 TACTGATACCTGTGTTTGAACAATGATGAACCGACACCTTAGTAGCTCAAGACTTGTTC 660
DB 601 TACTGATACCTGTGTTTGAACAATGATGAACCGACACCTTAGTAGCTCAAGACTTGTTC 660
QY 661 TCGGTTTACATGACAGAAATGCGCAGAGAGCATGCTCAGCTTGGAGCTTATGATCA 720
DB 661 TCGGTTTACATGACAGAAATGCGCAGAGAGCATGCTCAGCTTGGAGCTTATGATCA 720
QY 721 TCTTACTTACACAGATCTCTTCACTGAGTTCCAGTACTGTGACAGAGTACTGCAATTTC 780
DB 721 TCTTACTTACACAGATCTCTTCACTGAGTTCCAGTACTGTGACAGAGTACTGCAATTTC 780
QY 781 ACTGTGACAGAGTGTACCATCATGACCGGTGTGTTGTCATGTGAAAGTGTCAAGCT 840
DB 781 ACTGTGACAGAGTGTACCATCATGACCGGTGTGTTGTCATGTGAAAGTGTCAAGCT 840
QY 841 ATCTTGATACCGGTACGTCCAAAGCTGTGAGACCTAGACGCAATTCATTCAG 900
DB 841 ATCTTGATACCGGTACGTCCAAAGCTGTGAGACCTAGACGCAATTCATTCAG 900
QY 901 CAAGCTATTGAGCCCAACAAGAACCGTAGCGTAGTTGAATGATTCGACAACTT 960
DB 901 CAAGCTATTGAGCCCAACAAGAACCGTAGCGTAGTTGAATGATTCGACAACTT 960
QY 961 AGCTATACGCTTCAAGTGTGAGATCAACCGCAAGATGATACCACTGACCCCTTC 1020
DB 961 AGCTATACGCTTCAAGTGTGAGATCAACCGCAAGATGATACCACTGACCCCTTC 1020
QY 1021 GCCTATACCAAGCAGGATCAAGGTTCTGACAGTGTGATTCAGAGTGAACAATTC 1080
DB 1021 GCCTATACCAAGCAGGATCAAGGTTCTGACAGTGTGATTCAGAGTGAACAATTC 1080
QY 1081 CAGAAATGATCTTGGAGATGTTCATTCTGAGTACTACAGCGTCTTTGACAGGGCC 1140
DB 1081 CAGAAATGATCTTGGAGATGTTCATTCTGAGTACTACAGCGTCTTTGACAGGGCC 1140
QY 1141 AACCACTGTTGGGCTAGCTAAAGCAATCTGA 1173
DB 1141 AACCACTGTTGGGCTAGCTAAAGCAATCTGA 1173

RESULT 2
AAS00570
ID AAS00570 standard; DNA; 3957 BP.
XX
AC AAS00570;
XX
DT 14-MAY-2001 (first entry)
XX
DE Bovine phaseolin promoter pre-pro-chymosin-phaseolin terminator.
XX
KW Chymosin; transcription regulator; terminator sequence; soybean; corn;
KW pre-pro-chymosin; rape seed; sunflower; cotton; tobacco; alfalfa; wheat;
KW barley; oats; sorghum; Arabidopsis thaliana; potato; flax; linseed; rice;
KW safflower; oil palm; ground nut; Brazil nut; coconut; castor; coriander;
KW squash; jojoba; dei; phaseolin; promoter; terminator; mutant; French bean.
XX
OS Chimeric - Bos sp.
XX
OS Chimeric - Phaseolus vulgaris.
XX
FH Key Location/Qualifiers
FT 1..1553
FT /*tag= a
FT /note= "Phaseolin promoter"
FT CDS 1554..2726
FT /*tag= b
FT /product= "Bovine pre-pro-chymosin"
FT Terminator 2727..3957

```

```

FT      /*tag= C
XX      /note= "Phaseolin terminator"
XX      MO200114571-A1.
XX      01-MAR-2001.
XX      23-AUG-2000; 2000MO-CA00975.
XX      23-AUG-1999; 99US-0378696.
XX      (SEMB-) SEMBIOSYS GENETICS INC.
XX      Van Rooijen G, Keon RG, Boothe J, Shen Y;
XX      WPI, 2001-226621/23.
XX      P-PSDB; AAU00536.
XX      Producing chymosin in seeds of plants such as rice, flax, rape seed, by
XX      transforming plant cell with a nucleic acid encoding chymosin operably
XX      linked to transcription regulator and terminator sequences -
XX      Example 1; Fig 2; 56pp; English.
XX      The sequence represents a chimeric polynucleotide comprising a
XX      pre-pro-chymosin, a phaseolin promoter and a phaseolin terminator.
XX      Chymosin can be produced in a plant seed through introduction of a
XX      chimeric nucleic acid molecule, comprising a nucleic acid sequence
XX      encoding a chymosin polypeptide operatively linked to transcription
XX      regulator and terminator sequences, into a plant cell. The sequences are
XX      useful for producing plant seeds, in particular seeds of soybean, rape
XX      seed, sunflower, cotton, corn, tobacco, alfalfa, wheat, barley, oats,
XX      sorghum, Arabidopsis thaliana, potato, flax/linseed, safflower, oil palm,
XX      groundnut, Brazil nut, coconut, castor, coriander, squash, jojoba and
XX      rice.
XX      Sequence 3957 BP; 1263 A; 790 C; 609 G; 1295 T; 0 other;
XX
Query Match      100.0%; Score 1173; DB 22; Length 3957;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 ATGAATTCCTTAAGTCTTCCCTTTCTACGCTTTCCTTTGGTGGCAATACCTTGGT 60
DB      1554 ATGAATTCCTTAAGTCTTCCCTTTCTACGCTTTCCTTTGGTGGCAATACCTTGGT 1613
QY      61 GCTGTTACTCAGCTGCTGAGATCACCGGCAATTCCTTCTACAAAGTAAAGTCTCCGT 120
DB      1614 GCTGTTACTCAGCTGCTGAGATCACCGGCAATTCCTTCTACAAAGTAAAGTCTCCGT 1673
QY      121 AAGGCGCTGAAGGAATGAGCTTCTAGAAAGCTTCTTGCAAGAAACAAGATAGGCATC 180
DB      1674 AAGGCGCTGAAGGAATGAGCTTCTAGAAAGCTTCTTGCAAGAAACAAGATAGGCATC 1733
QY      181 AGCAGCAAGTACTCCGGCTCGTGAAGTGTGCTAGCGCACTTACCAACTACCTTGTAT 240
DB      1734 AGCAGCAAGTACTCCGGCTCGTGAAGTGTGCTAGCGCACTTACCAACTACCTTGTAT 1793
QY      241 AGTCAATACCTTTGGGAAGATCTACTCTGGAACCCCGCTCAAGAGTTCAACGGTCTCTT 300
DB      1794 AGTCAATACCTTTGGGAAGATCTACTCTGGAACCCCGCTCAAGAGTTCAACGGTCTCTT 1853
QY      301 GATACCTGTTCTCTGACCTTCTGGGTTCCCTCTATCTCTGTAAGAAGAAAGCCCTGCAAG 360
DB      1854 GATACCTGTTCTCTGACCTTCTGGGTTCCCTCTATCTCTGTAAGAAGAAAGCCCTGCAAG 1913
QY      361 AACCAACCAAGATTGATCCGAGAAAGTGTCTCAACTTTCAGAACTTAGGCAAAACCTTGG 420
DB      1914 AACCAACCAAGATTGATCCGAGAAAGTGTCTCAACTTTCAGAACTTAGGCAAAACCTTGG 1973
QY      421 TCTATACCTAGAGTACAGTATGATGCAAGAAATCTTAGGCTATGATACCTGCTACTGTC 480
DB      1974 TCTATACCTAGAGTACAGTATGATGCAAGAAATCTTAGGCTATGATACCTGCTACTGTC 2033

```

```

QY      481 TCCAACATTTGGACATTCACACAGACAGTATGAGACTTTAGCACCCAGAAACAGGTATGTC 540
DB      2034 TCCAACATTTGGACATTCACACAGACAGTATGAGACTTTAGCACCCAGAAACAGGTATGTC 2093
QY      541 TTCACCTATGCAAGAAATTTGATGTCATCTTGTATGATGATACCCATTCGCTGCGCTCAGAG 600
DB      2094 TTCACCTATGCAAGAAATTTGATGTCATCTTGTATGATGATACCCATTCGCTGCGCTCAGAG 2153
QY      601 TACTCGATACCTGTTGTTGACAACTGATGAAACCGACACTAGTATGCTCAAGACTTGTTC 660
DB      2154 TACTCGATACCTGTTGTTGACAACTGATGAAACCGACACTAGTATGCTCAAGACTTGTTC 2213
QY      661 TCGGTTTACATGACAGGAATGCGCAGAGAGACATGCTCAAGCTTGGAGCTATTGATCCA 720
DB      2214 TCGGTTTACATGACAGGAATGCGCAGAGAGACATGCTCAAGCTTGGAGCTATTGATCCA 2273
QY      721 TCTTACTACACAGGATCTCTTCACTGCGGTTCCAGTCACTGTGACGAGTACTGCGCAATTC 780
DB      2274 TCTTACTACACAGGATCTCTTCACTGCGGTTCCAGTCACTGTGACGAGTACTGCGCAATTC 2333
QY      781 ACTGTGACAGTGTCCACCATGACAGCGGTGTGTTGTTGATGTAAGTGTCAAGCT 840
DB      2334 ACTGTGACAGTGTCCACCATGACAGCGGTGTGTTGTTGATGTAAGTGTCAAGCT 2393
QY      841 ATCTTGATACCGGTAAGTCCCAAGCTGTCGAGACCTTACAGCACTTCAACATTCAG 900
DB      2394 ATCTTGATACCGGTAAGTCCCAAGCTGTCGAGACCTTACAGCACTTCAACATTCAG 2453
QY      901 CAAGCTATTTGAGCCACACAGAACCACTAGCGTGTGATTTGAATATGATTTGGACAACTT 960
DB      2454 CAAGCTATTTGAGCCACACAGAACCACTAGCGTGTGATTTGAATATGATTTGGACAACTT 2513
QY      961 AGCTACATGCTTACAGTGTCTTTGATGATCAACGGCAAGATGTACCCACTGACCCCTTC 1020
DB      2514 AGCTACATGCTTACAGTGTCTTTGATGATCAACGGCAAGATGTACCCACTGACCCCTTC 2573
QY      1021 GCTATACACAGCCAGATCAAGGGTTCGACACAGTGTATCCAGAGTGAAGAACCTTCC 1080
DB      2574 GCTATACACAGCCAGATCAAGGGTTCGACACAGTGTATCCAGAGTGAAGAACCTTCC 2633
QY      1081 CAGAAATGATCTTTGGAGATGTGTCTATTCGTGATGATCTACAGCGTCTTTGACAGGGCC 1140
DB      2634 CAGAAATGATCTTTGGAGATGTGTCTATTCGTGATGATCTACAGCGTCTTTGACAGGGCC 2693
QY      1141 AACCACTCGTTGGGCTAGCTTAAGCAATCTGA 1173
DB      2694 AACCACTCGTTGGGCTAGCTTAAGCAATCTGA 2726

```

RESULT 3
 AAN30063
 ID AAN30063 standard; cDNA; 1098 BP.
 AC AAN30063;
 AC AAN30063;
 DT 14-JUN-1992 (first entry)
 DE Sequence of prorennin cDNA in PCR 10001.
 KW Remnin; renin; enzyme; protease; ss.
 OS Bos taurus.
 XX Key Location/Qualifiers
 FH CDS 1..1098
 FT /*tag= a
 PN EP73029-A.
 PD 02-MAR-1993.
 PF 19-AUG-1982; 82EP-0107601.

```
XX 24-AUG-1981; 81JP-0131631.
XX (BEPF/) BEPPU T.
XX Bepu T, Uozumi T, Nishimori K;
XX WPI: 1983-22976K/10.
XX P-PSDB; AAP90603.
XX plasmid contg. calf pro:rennin DNA - and transformed
XX microorganisms
XX
XX Example; Page 20-23; 32pp; English.
XX
XX The inventors claim recombinant plasmids contg. the cDNA of calf
XX prorennin. Specified plasmids are PCR 10001 and PCR2001 (contg. the
XX whole sequence plus the lac promoter region). Also new are
XX microorganisms transformed with the plasmids, esp. E. coli CR1 (ATCC
XX 391710) contg. plasmid PCR2001.
XX
XX Sequence 1098 BP; 252 A; 326 C; 301 G; 219 T; 0 other;
XX
XX Query Match 4.2%; Score 49; DB 4; Length 1098;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-14;
XX Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 656 TGTTCGCGTTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCGT 704
XX 581 TGTTCGCGTTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCGT 629
XX
XX RESULT 4
XX AAQ04683
XX ID AAQ04683 standard; DNA; 1098 BP.
XX
XX AAQ04683;
XX
XX 25-MAR-2003 (updated)
XX 05-OCT-1990 (first entry)
XX
XX Sequence encoding calf pro-rennin.
XX
XX Pro-rennin; ds.
XX
XX Bos taurus.
XX
XX JP02109984-A.
XX
XX 23-APR-1990.
XX
XX 28-OCT-1981; 81JP-0302176.
XX
XX 28-OCT-1981; 81JP-0302176.
XX
XX (BEPF/) BEPPU T.
XX
XX WPI: 1990-168358/22.
XX P-PSDB; AAR05080.
XX
XX Complex plasmid and microbe - contains calf pro-rennin cDNA.
XX
XX Disclosure; 32; 13pp; Japanese.
XX
XX Protein product may be expressed in E.coli expression system from
XX plasmid pBR322.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX (Updated on 25-MAR-2003 to correct PR field.)
XX
XX Sequence 1098 BP; 252 A; 327 C; 300 G; 219 T; 0 other;
XX
XX Query Match 4.2%; Score 49; DB 11; Length 1098;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-14;
```

```
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 656 TGTTCGCGTTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCGT 704
XX 581 TGTTCGCGTTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCGT 629
XX
XX RESULT 5
XX AAN40295
XX ID AAN40295 standard; mRNA; 1175 BP.
XX
XX AAN40295;
XX
XX 04-FEB-1992 (first entry)
XX
XX Sequence encoding a polypeptide displaying milk clotting activity.
XX
XX Cheese-making; recombinant protein; rennet substitute; milk clot; ss.
XX
XX Bos taurus.
XX
XX Key Location/Qualifiers
XX sig_peptide 21..69
XX mat_peptide 70..1166
XX /tag= b
XX
XX EP123928-A.
XX
XX 07-NOV-1984.
XX
XX 30-MAR-1984; 84BP-0103551.
XX
XX 31-MAR-1983; 83US-0480860.
XX
XX (CODO-) CODON GENETIC ENG.
XX
XX Cashion LM, McCaman MT, Rice CW, Sias SR;
XX
XX WPI: 1984-277277/45.
XX P-PSDB; AAP40559.
XX
XX Recombinant DNA coding for milk clotting polypeptide - which is
XX expressed in transformed bacteria
XX
XX Claim 6; Fig 2; 39pp; English.
XX
XX Also claimed is E.coli JM83/PLC7 (ATCC 39325) which is transformed
XX with PLC7 contg. the prorennin derived sequence fused in phase with
XX B-galactosidase. The PLC7 prorennin expression plasmid includes
XX sequences which code for both the pseudorennin and mature rennin
XX cleavage sites between Aas 28-29 and Aas 42-43, respectively.
XX
XX Sequence 1175 BP; 263 A; 355 C; 319 G; 238 U; 0 other;
XX
XX Query Match 4.2%; Score 49; DB 5; Length 1175;
XX Best Local Similarity 75.5%; Pred. No. 1.5e-14;
XX Matches 37; Conservative 12; Mismatches 0; Indels 0; Gaps 0;
XX
XX 656 TGTTCGCGTTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCGT 704
XX 649 UGUUCUGGUUDACUAGACAGGAUUGCCAGAGAGCAUGCUCACGCU 697
XX
XX RESULT 6
XX AAQ20949
XX ID AAQ20949 standard; cDNA; 1175 BP.
XX
XX AAQ20949;
XX
XX 13-MAY-1992 (first entry)
XX
XX Prochymosin (prorennin) gene from calf.
```


XX	Rennet; zymogen; autocatalytic activation; pseudorennin; milk;
KM	clotting activity; ss.
OS	Bos taurus.
XX	
FH	Key
FT	CDS
XX	
PN	US082775-A.
XX	
PB	21-JAN-1992.
XX	
PF	28-OCT-1988; 88US-0263927.
PR	28-OCT-1988; 88US-0263927.
PR	11-MAY-1984; 84US-0609495.
PR	12-DEC-1986; 86US-0940199.
PR	31-MAR-1983; 83US-0480860.
XX	28-APR-1986; 86US-0856700.
PA	(BERL-) BERLEX LABS INC.
XX	
P1	McCaman WT, King JF;
DR	WPI; 1992-049149/06.
DR	P-PsDB; AAR20730.
PT	Isolating heterologous polypeptide from bacterial inclusion
PT	bodies - by lysing cells, extrn. with nonionic detergent and sepy.
PS	insoluble polypeptide
XX	
PS	Disclosure; Fig 6; 21pp; English.
CC	The prorennin (prochymosin) sequence was obt'd.from mRNA from
CC	unweaned calf's stomach (abomasum). To identify rennin-specific
CC	mRNA species, hybridisation with rennin specific probes was performed
CC	using probes derived from the published amino acid sequence of
CC	prorennin (Poltmann et al., J. Biol. Chem. 254, 8447-8456 (1979))
CC	(see AAQ23291.2). Reverse transcriptase was used to transcribe mRNA
CC	into a cDNA copy which was inserted into plasmid pBR322 and cloned
CC	into E. coli strain K-12 NM 294 (ATCC 31446). The prorennin
CC	specific clones were used to screen for recombinant plasmids.
CC	Only two clones 5G3 and 15C5 were found to contain the whole prorennin
CC	sequence, with clone 5G3 having its complete coding sequence (shown
CC	here). The sequence was the same as the published sequence except
CC	for two silent mutations in codons 274 and 336 respectively, and a
CC	single mutation in codon 302 which converts the sequence to the B form
CC	of rennin. The product of prorennin, rennin is an active component of
CC	rennet which is used to clot milk in the process of making cheese.
CC	See also AAQ20950.
SQ	Sequence 1175 BP; 261 A; 355 C; 319 G; 240 U; 0 other;
XX	
Query Match	4.2%; Score 49; DB 13; Length 1175;
Best Local Similarity	75.5%; Pred. No. 1.5e-14;
Matches 37; Conservative 12; Mismatches 0; Indels 0; Gaps 0	
Dy	656 TGGTTCGCAGTTACATGACAGAAATGACCAGAGAAGCATGCTCACGCT 704 ::: :::: : : : : : :
Dd	649 UGUUCUGCGUUAUCAUGACAAGAUUGGCCAGGAAGCAUGCUGUACGCU 697 ::: :::: : : : : : :
RESULT 7	
ID	AAQ14051 standard; DNA; 1210 BP.
AC	AAQ14051;
XX	
DT	25-MAR-2003 (updated)

DT	06-JAN-1992	(first entry)	
XX			
DE	Remnin gene.		
KW	Prorennin; alpha-S1-casein gene; insulin-like growth factor I; IGF-I;		
XX	mammary gland; ss.		
OS	Synthetic.		
XX			
PN	EP451823-A.		
XX			
PD	16-OCT-1991.		
XX			
PF	10-APR-1991; 91EP-0105702.		
XX			
PR	19-APR-1990; 90DE-4012526.		
PR	11-APR-1990; 90DE-4011751.		
XX			
PA	(CONE) CONSORTIUM ELEKTROCHEM IND.		
XX			
P1	Hartl P, Brem G;		
DR	WPI; 1991-304858/42.		
XX			
PT	Recombinant DNA constructs for expressing protein in milk -		
PT	contg. specific mammary gland transcription control region and		
PT	signal sequence, providing high yield and easy prod. recovery		
XX			
PS	Disclosure; Page 21; 41pp; German.		
XX			
CC	The remnin gene was isolated using the probe represented in AA014775.		
CC	It was used as heterologous peptide/protein together with parts of the		
CC	alpha-S1-casein gene in the prodn. of DNA constructs. The heterologous		
CC	peptide or protein may also be human insulin-like growth factor I.		
CC	The constructs provide high yields of the protein with simple recovery		
CC	from the milk. Activation of the gene occurs only in the mammary gland.		
CC	See also AA014050, AA014774-77.		
CC	(Updated on 25-MAR-2003 to correct PA field.)		
XX			
SO	Sequence 1210 BP; 278 A; 356 C; 324 G; 252 T; 0 other;		
	Query Match	4.2%; Score 49; DB 12; Length 1210;	
	Best Local Similarity	100.0%; Pzed. No. 1.5e-14;	
	Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	656 TGGTCTCGGTTTACATGACGAGAAATGCGCAGAGAGCATGCTCAGCGT 704		
DB	643 TGTTCCTGGTTTACATGACGAGAAATGCGCAGAGAGCATGCTCAGCGT 691		
	RESULT 8		
ID	AA03006		
ID	AA03006 standard; DNA; 1240 BP.		
AC	AA03006;		
XX			
DT	13-JUN-1996 (first entry)		
XX			
DE	Chymosin open reading frame.		
XX			
XX	a1.3; albino mutant; light-regulated; Neurospora; bread mould;		
KW	heterologous gene; expression; control; chymosin; ss.		
XX			
OS	Mammalian sp.		
XX			
PN	W09530739-A1.		
PD	16-NOV-1995.		
XX			
PF	09-MAY-1995; 95WO-US05716.		
XX			
PR	10-MAY-1994; 94US-0240372.		
XX			

PA (UYHA-) UNIV HAWAII.
 XX Kato EK, Stuart WD;
 PI MPI, 1995-404108/51.
 DR Nucleic acid for expression of heterologous protein - contains
 XX albedo promoter for light induced expression in filamentous fungi
 PT Example 2; Fig 7, 26pp; English.
 PS
 XX The DNA is that of a mammalian gene (open reading frame) encoding
 CC chymosin. The gene was placed in operable linkage with the al-3
 CC promoter (see AAT03005). The al-3 gene controls the production of
 CC geranyl geranyl pyrophosphatase (GGPP) synthetase. GGPP is a precursor
 CC for carotenoids and xanthophylls. It has been shown that exposure to
 CC light increases the transcription level of GGPP synthetase 15-45 fold.
 CC This can be used to regulate the expression of genes encoding
 CC heterologous proteins, e.g. chymosin, in recombinant production systems.
 CC Use of a light-regulated promoter is a simple and effective way to
 CC control expression and allows timing to be adapted to the physiological
 CC status of the host.
 CC
 SQ Sequence 1240 BP; 274 A; 374 C; 339 G; 253 T; 0 other;
 Query Match 4.2%; Score 49; DB 16; Length 1240;
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 656 TGTTCGCGTTTACATGACAGGAATGCGCAGAGAGCATGCTCAGCGT 704
 Db 700 TGTTCGCGTTTACATGACAGGAATGCGCAGAGAGCATGCTCAGCGT 748
 RESULT 9
 AAN40055
 ID AAN40055 standard; DNA; 1278 BP.
 AC AAN40055;
 XX
 DT 02-FEB-1992 (first entry)
 XX
 DE Sequence of prochymosin gene.
 XX
 KW Prochymosin expression vector; E.coli trp operon; chymosin; ss.
 XX
 OS Bos taurus.
 XX
 FH Key Location/Qualifiers
 FT CDS 3..1130
 FT /*tag= a
 FT polyA_signal 1245..1250
 FT /*tag= b
 XX
 PN EP121775-A.
 XX
 PD 17-OCT-1984.
 XX
 PF 07-MAR-1984; 84BP-0102451.
 XX
 PR 09-MAR-1983; 83JP-0038439.
 XX
 PA (BEPP/) BEPPU T.
 XX
 PI Beppu T, Uozumi T, Nishimori K, Shimizu N, Kawaguchi Y;
 PI Hataka M;
 XX WPI; 1984-258001/42.
 DR P-PSDB; AAP40078.
 XX
 PT Expression plasmid comprising prochymosin gene and vector -
 PT useful for transforming Escherichia coli for prochymosin prodn.

XX Disclosure; Fig 1; 59pp; English.
 XX
 CC The inventors claim the prochymosin gene comprising a nucleotide
 CC sequence from (a) the 1st codon (GCT) to the 365th codon (ATC), or
 CC (b) the 5th codon (CGG) to the 365th codon (ATC); and recombinant
 CC plasmids harboured by Escherichia coli strains deposited as FERM BP-
 CC 262, -263 and -264. Any portion of the nucleotide sequence as
 CC described in AAN40055 can be used. Also claimed is a vector derived
 CC from plasmid pBR322. Typically plasmid pBR501 is used. from pOCT 2.
 CC The transcriptional direction of pOCT 3 is opposite to that of
 CC pOCT 2; it is clockwise in pOCT 2 (5' to 3') whereas counter
 CC clockwise in pOCT 3.
 CC
 SQ Sequence 1278 BP; 309 A; 384 C; 326 G; 259 T; 0 other;
 Query Match 4.2%; Score 49; DB 5; Length 1278;
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 656 TGTTCGCGTTTACATGACAGGAATGCGCAGAGAGCATGCTCAGCGT 704
 Db 613 TGTTCGCGTTTACATGACAGGAATGCGCAGAGAGCATGCTCAGCGT 661
 RESULT 10
 AAN30022
 ID AAN30022 standard; DNA; 1289 BP.
 AC AAN30022;
 XX
 DT 25-MAR-2003 (updated)
 DT 25-APR-1992 (first entry)
 XX
 DE Sequence of veal chymosin gene.
 XX
 KW Protolytic enzyme; zymogen; rennin; chymosin; cheese making;
 KW microbial vector; ss.
 XX
 OS Bos taurus.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1140
 FT /*tag= a
 XX
 PN BE897201-A.
 XX
 PD 03-NOV-1983.
 XX
 PF 30-JUN-1983; 83BE-0017731.
 XX
 PR 13-APR-1983; 83US-0484539.
 PR 01-JUL-1982; 82US-0394433.
 XX
 PA (GENEX) GENEX CORP.
 XX
 DR WPI; 1983-820813/47.
 DR P-PSDB; AAP30013.
 XX
 PT Isolated chymosin or rennin and prochymosin genes - plasmid(s)
 PT which replicate in prokaryotic organisms, esp. Escherichia coli,
 PT and organisms used for chymosin biosynthesis
 XX
 PS Claim 7; Page 33-36; 43pp; French.
 XX
 CC The inventors claim isolated chymosin (rennin) and prochymosin genes
 CC from calves, and plasmids contg. the genes which are capable of
 CC replicating in a prokaryotic organism. The prokaryotic organism is
 CC pref. an Escherichia species, esp. E. coli p Gx 1225 (NRRL B-15061).
 CC The microorganisms transformed by the plasmid are also claimed.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC
 SQ Sequence 1289 BP; 310 A; 391 C; 327 G; 260 T; 1 other;

Query Match 4.2%; Score 49; DB 4; Length 1289;
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 656 TGTTCGCGTTTACATGACAGAGATGGCCAGAGAGCATGCTCAGCT 704
 DB 623 TGTTCGCGTTTACATGACAGAGATGGCCAGAGAGCATGCTCAGCT 671

RESULT 11

AA91157
 ID AAN91157 standard; DNA; 1291 BP.

AC AAN91157;

DT 10-MAR-2003 (updated)
 DT 07-JUN-1990 (first entry)

DE Cloned sequence of (pro)chymosin.

KW Lactic acid bacteria; cheese; Streptococcus cremoris SK112; chymosin;
 KM proteinase; pSK112; ss.

OS Unidentified.

Key Location/Qualifiers
 FT precursor_RNA 29.1123
 FT /*tag= a
 FT /product=;prochymosin

PN NL8701378-A.

PD 02-JAN-1989.

PF 12-JUN-1987; 87NL-0001378.

PR 12-JUN-1987; 87NL-0001378.

PS (NEZU-) NEDERL INS ZUIVELON.

PI Simons AFM; De Vos WM;

DR WPI; 1989-030097/04.

DR P-PSDB; AAP94144.

FT DNA fragment having region specific for lactic acid bacteria -
 FT is contained in plasmid in microorganism used in prodn. of
 PT protein and food prodn. eg cheese.

PS Disclosure; fig 2; 43pp; Dutch.

XX The DNA encoding prochymosin can be cloned into a plasmid (esp. from
 CC S. cremoris SK112) and used to produce large amts of the protein by
 CC recombinant DNA techniques. This could overcome the shortage of
 CC prochymosin due to shortage of calf stomachs and increasing cheese
 CC prodn. Prochymosin is also used in prodn. of yoghurt, butter and
 CC buttermilk. See also AAN91158-N91160.
 CC (Updated on 10-MAR-2003 to add missing OS field.)

CC Sequence 1291 BP; 323 A; 385 C; 326 G; 257 T; 0 other;

Query Match 4.2%; Score 49; DB 10; Length 1291;
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 656 TGTTCGCGTTTACATGACAGAGATGGCCAGAGAGCATGCTCAGCT 704
 DB 609 TGTTCGCGTTTACATGACAGAGATGGCCAGAGAGCATGCTCAGCT 657

RESULT 12

AAN30049

ID AAN30049 standard; cDNA; 1314 BP.

XX AAN30049;

DT 14-JUN-1992 (first entry)

DE cDNA sequence corresponding to one of the allelic forms (B) of
 DE bovine preprochymosin.

XX Chymosin; enzyme; rennet; cheese; ss.

OS Bos taurus.

Key Location/Qualifiers
 FT sig_peptide 24..40
 FT /*tag= a
 FT mat_peptide 41..1169
 FT /*tag= b

PN EP77109-A.

PD 20-APR-1983.

PF 13-OCT-1982; 82EP-0201272.

PR 14-OCT-1981; 81GB-0031004.

PS (UNIL) UNILEVER NV.

PI Maat J, Verrips CT, Ledebor AM, Edens L;

DR WPI; 1983-39656K/17.

DR P-PSDB; AAP30086.

FT DNA molecules comprising genes for preprochymosin - used to
 PT transform microorganisms to give strain producing the
 PT prepro-enzyme and its allelic and maturation forms

PS Claim 2; Fig 1; 53pp; English.

XX Preprochymosin is an intermediate (via prochymosin and
 CC pseudochymosin) for the enzyme chymosin, which is the essential milk-
 CC clotting component of rennet and is used in cheese manufacture.
 CC AAN30049 corresp. to mRNA isolated from the fourth stomach of a
 CC preterminant calf (abomasum, Frisian cow).

XX Sequence 1314 BP; 309 A; 398 C; 338 G; 269 T; 0 other;

Query Match 4.2%; Score 49; DB 4; Length 1314;
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 656 TGTTCGCGTTTACATGACAGAGATGGCCAGAGAGCATGCTCAGCT 704
 DB 652 TGTTCGCGTTTACATGACAGAGATGGCCAGAGAGCATGCTCAGCT 700

RESULT 13

AAN20043
 ID AAN20043 standard; DNA; 1460 BP.

AC AAN20043;

DT 16-DEC-1992 (first entry)

DE Pre-prorennin-A gene DNA sequence.

XX Pre-pro-remnin; rennin; prorennin; enzyme; EC-3.4.23.4; chymosin;
 KW protease; milk-clotting enzyme; ss.

OS Bos taurus.

Key Location/Qualifiers

FT CDS 205..1350
FT /**tag= a
FT /label=
FT /note="pre-prorennin-A gene"
XX
XX GB2091271-A.
XX
XX 28-JUL-1982.
XX
XX 15-JAN-1982; 82GB-0001120.
XX
XX 01-DEC-1981; 81US-0325481.
XX 16-JAN-1981; 81US-0225717.
XX
XX (COLB) COLLABORATIVE RES INC.
XX
XX Alford BL, Mao J, Moir DT;
XX
XX WPI; 1982-62028E/30 (62028B).
XX P-PSDB; AAP20038.
XX
XX Transformed cells producing rennin and its precursors - contg.
XX appropriate recombinant DNA material
XX
XX Disclosure; Table 1; 39pp; English.
XX
XX Bases 1-204 and 1351 to 1460 are attached to the pre-prorennin but
XX can be removed and are not essential to use of the gene in
XX expression. The gene may be ligated into plasmid pCGE21 and
XX expressed in E. coli. The resulting expressed enzyme is a well
XX known milk-clotting enzyme used in cheese-making.
XX
XX Sequence 1460 BP; 329 A; 440 C; 397 G; 294 T; 0 other;
SQ
Query Match 4.2%; Score 49; DB 3; Length 1460;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 656 TGTTCCTCGTTTACATGACAGGAATGCGCAGAGACATGCTCAGCT 704
DB 833 TGTTCCTCGTTTACATGACAGGAATGCGCAGAGACATGCTCAGCT 881
RESULT 14
AAN40180
ID AAN40180 standard; CDNA; 1460 BP.
XX
XX AAN40180;
XX
XX 25-MAR-2003 (updated)
DT 25-JAN-1992 (first entry)
XX
XX Sequence of recombinant CGF4 carrying the rennin coding sequence.
XX
XX Yeast expression vector; GAL1 promoter; Saccharomyces cerevisiae;
XX es.
XX Bos taurus.
XX
XX OS
XX Key Location/Qualifiers
FH CDS 205..1350
FT /**tag= a
XX
XX GB2137208-A.
XX
XX 03-OCT-1984.
XX
XX 28-FEB-1984; 84GB-0005129.
XX
XX 28-FEB-1983; 83US-0470911.
XX
XX (COLB) COLLABORATIVE RES INC.
XX

PI Botstein D, Davis RW, Fink GR, Tauntonrig A, Knowlton RG, Mao JI;
PI Moir DT, Goff CG;
XX
XX WPI; 1984-245517/40.
XX
XX P-PSDB; AAP40218.
XX
XX DNA segment contg. GAL1 promoter linked to gene - useful for
XX direction of expression of the gene in yeast cell
XX
XX Example; Table 4, Page 21-23; 35pp; English.
XX
XX The inventors claim a DNA segment contg. GAL1 promoter linked to
XX gene - useful for direction of expression of the gene in yeast
XX cell.
XX The recombinant material carrying a GAL1 promoter of the yeast
XX galactokinase gene may be used in expressing a desired protein, esp.
XX bovine growth hormone, interferon, prorennin or preprorennin, in the
XX yeast cell. Strains of Saccharomyces cerevisiae producing the
XX polypeptides are produced. Yeast strains deposited as ATCC 20643,
XX 20661, 20662 and 20663, strain designations CGY 196, 457, 461 and
XX 528, resp. are new.
XX (updated on 25-MAR-2003 to correct 'P' field.)
XX (updated on 25-MAR-2003 to correct 'I' field.)
XX
XX Sequence 1460 BP; 328 A; 440 C; 399 G; 293 T; 0 other;
SQ
Query Match 4.2%; Score 49; DB 5; Length 1460;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 656 TGTTCCTCGTTTACATGACAGGAATGCGCAGAGACATGCTCAGCT 704
DB 833 TGTTCCTCGTTTACATGACAGGAATGCGCAGAGACATGCTCAGCT 881
RESULT 15
AAN91188
ID AAN91188 standard; DNA; 2727 BP.
XX
XX AAN91188;
XX
XX 25-MAR-2003 (updated)
DT 15-JUN-1990 (first entry)
XX
XX BamHI/SalI insert of Kluyveromyces plasmid pAB309.
XX
XX Kluyveromyces; pDM100PC; chymosin; tissue plasminogen activator.
XX
XX Kluyveromyces lactis.
XX
XX OS
XX Key Location/Qualifiers
FH CDS 409..1781
FT /**tag= a
XX
XX EP301670-A.
XX
XX 01-FEB-1989.
XX
XX 28-JUL-1988; 88EP-0201632.
XX
XX 28-JUL-1987; 87US-0078539.
XX
XX (KONN) GIST-BROCADES NV.
XX
XX Vandenberg JA, Vanooeyen AJJ, Rietveld K;
XX
XX WPI; 1989-033565/05.
XX P-PSDB; AAP94376.
XX
XX Kluyveromyces host cells for producing polypeptide(s) -
XX used for highly efficient prodn. of eg chymosin tissue
XX plasminogen activator or human serum albumin.
XX
XX Disclosure; Page 7; 56pp; English.
PS

XX BamHI/SalI insert, incorporated into plasmids pAB309 with
CC promoters, terminators and a G418 resistance marker fused to a ADH1
CC promoter from *S.cerevisiae*.
CC The Kluyveromyces expression systems provide highly efficient secretion
CC and processing of a wide variety of proteins.
CC Sequences identical to those published in EP901669.
CC (Updated on 25-MAR-2003 to correct PD field.)
CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 2727 BP; 752 A; 654 C; 572 G; 749 T; 0 other;
Query Match 4.2%; Score 49; DB 10; Length 2727;
Best Local Similarity 100.0%; Pred.No. 1.5e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 656 TGTTCCTCGGTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCGCT 704
|||||
Db 1268 TGTTCCTCGGTTACATGACAGGAATGGCCAGAGAGCATGCTCAGCGCT 1316
|||||

Search completed: February 5, 2004, 06:07:28
Job time : 396 secs

THIS PAGE BLANK (USPTO)

RESULT 2
US-08-846-021A-6
; Sequence 6, Application US/08846021A
; Patent No. 5948682
; GENERAL INFORMATION:
; APPLICANT: MOLONEY, Maurice M.
; TITLE OF INVENTION: Preparation of Heterologous Proteins on
; TITLE OF INVENTION: Oil Bodies
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERSKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/846.021A
; FILING DATE: April 25, 1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 9369-039
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 850..1206
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1444..2729
; US-08-846-021A-6
Query Match 4.2%; Score 49; DB 2; Length 2733;
Best Local Similarity 100.0%; Pred. No. 9.5e-16;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 656 TGTTCGCGTTTACATGACAGATGCGCCAGAGAGACATGCTCAGCCT 704
Db 2210 TGTTCGCGTTTACATGACAGATGCGCCAGAGAGACATGCTCAGCCT 2258
RESULT 3
5217891-14
; Patent No. 5217891
; APPLICANT: BRAKE, ANTHONY J.; VAN DEN BERG, JOHAN A.
; TITLE OF INVENTION: DNA CONSTRUCTS CONTAINING A KLUYVEROMYCES
; A FACTOR LEADER SEQUENCE FOR DIRECTING SECRETION OF HETEROLOGOUS
; POLYPEPTIDES
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/507.398
; FILING DATE: 09-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 78,551
; FILING DATE: 28-JUL-1987
; SEQ ID NO: 14:
; LENGTH: 2732
5217891-14

Query Match 3.2%; Score 38; DB 6; Length 2732;
Best Local Similarity 100.0%; Pred. No. 5.2e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 778 TTCACGTGACAGTGTACATCAGCGGTGTGTGT 815
Db 1391 TTCACGTGACAGTGTACATCAGCGGTGTGTGT 1428
RESULT 4
5200327-6
; Patent No. 5200327
; APPLICANT: GARVIN, ROBERT T.; MALEK, LAWRENCE T.
; TITLE OF INVENTION: EXPRESSION SYSTEM FOR THE SECRETION OF
; BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY STIMULATING
; FACTOR (GM-CSF) AND OTHER HETEROLOGOUS PROTEINS FROM
; STREPTOMYCES
; NUMBER OF SEQUENCES: 24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/224.568
; FILING DATE: 26-JUL-1988
; SEQ ID NO: 6:
; LENGTH: 828
5200327-6
Query Match 1.7%; Score 20; DB 6; Length 828;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 358 AAGACCACCAAGATTGCA 377
Db 795 AAGACCACCAAGATTGCA 814
RESULT 5
US-08-118-193-15
; Sequence 15, Application US/08118193
; Patent No. 564163
; GENERAL INFORMATION:
; APPLICANT: GARVIN, Robert T.
; APPLICANT: MALEK, Lawrence T.
; TITLE OF INVENTION: OF BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY
; STIMULATING FACTOR (GM-CSF) AND OTHER HETEROLOGOUS
; TITLE OF INVENTION: PROTEINS FROM STREPTOMYCES
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318.193
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.314
; FILING DATE:
; APPLICATION NUMBER: US 07/224.568
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 18740/116 CACO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid:
DESCRIPTION: Synthetic DNA oligonucleotide
FEATURE:
NAME/KEY: CDS
LOCATION: 399..830
US-08-318-193-15

Query Match 1.7%; Score 20; DB 1; Length 832;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 AAGAACCCCAAGATTGCA 377
|||||
DB 795 AAGAACCCCAAGATTGCA 814

RESULT 6
US-08-318-193-13
Sequence 13, Application US/08318193
Patent No. 5641663
GENERAL INFORMATION:
APPLICANT: GARVIN, Robert T.
TITLE OF INVENTION: AN EXPRESSION SYSTEM FOR THE SECRETION
TITLE OF INVENTION: OF BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY
TITLE OF INVENTION: STIMULATING FACTOR (GM-CSF) AND OTHER HETEROLOGOUS
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,193
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,314
FILING DATE:
APPLICATION NUMBER: US 07/224,568
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/116 CACO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)883-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 838 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid:
DESCRIPTION: Synthetic DNA oligonucleotide
FEATURE:

NAME/KEY: CDS
LOCATION: 399..836
US-08-318-193-13

Query Match 1.7%; Score 20; DB 1; Length 838;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 AAGAACCCCAAGATTGCA 377
|||||
DB 801 AAGAACCCCAAGATTGCA 820

RESULT 7
5200327-5
Patent No. 5200327
APPLICANT: GARVIN, ROBERT T.; MALEK, LAWRENCE T.
TITLE OF INVENTION: EXPRESSION SYSTEM FOR THE SECRETION OF
BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY STIMULATING
FACTOR (GM-CSF) AND OTHER HETEROLOGOUS PROTEINS FROM
STREPTOMYCES
NUMBER OF SEQUENCES: 24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/224,568
FILING DATE: 26-JUL-1988
SEQ ID NO: 5;
LENGTH: 846
5200327-5

Query Match 1.7%; Score 20; DB 6; Length 846;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 AAGAACCCCAAGATTGCA 377
|||||
DB 801 AAGAACCCCAAGATTGCA 820

RESULT 8
US-09-149-476-263/c
Sequence 263, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048, 964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057, 650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056, 884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057, 669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049, 610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061, 060
EARLIER FILING DATE: 1997-10-02

Query Match 1.6%; Score 19; DB 4; Length 689;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1068 TGAGAACCATCCCGAGAA 1086
Db 323 TGAGAACCATCCCGAGAA 305

RESULT 9
US-09-149-476-110/c
Sequence 110, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002p1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040, 162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040, 333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038, 621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040, 626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040, 334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040, 336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040, 163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047, 600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 618
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047, 503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043, 580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048, 974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056, 886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 637

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 1.6%; Score 19; DB 4; Length 1751;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1068 TGAGAACCATTCACGAAA 1086
Db 1101 TGAGAACCATTCACGAAA 1083

RESULT 10
US-09-738-884-3/C
Sequence 3, Application US/09738884
Patent No. 6391606
GENERAL INFORMATION:
APPLICANT: GUEGLER, Karl et al
TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,
NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
FILE REFERENCE: CL000849
CURRENT APPLICATION NUMBER: US/09/738,884
CURRENT FILING DATE: 2000-12-18
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 13953
TYPE: DNA
ORGANISM: Human
US-09-738-884-3

Query Match 1.6%; Score 19; DB 4; Length 13953;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 GCATTCCTCTCTCAAG 107
Db 10698 GCATTCCTCTCTCAAG 10680

RESULT 11
US-09-107-532A-1731
Sequence 1731, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
City: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 1731:
SEQUENCE CHARACTERISTICS:
LENGTH: 1746 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1..1746
SEQUENCE DESCRIPTION: SEQ ID NO: 1731:
US-09-107-532A-1731

Query Match 1.5%; Score 18; DB 4; Length 1746;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 122 AGGCGCTGAGGACATG 139
Db 509 AGGCGCTGAGGACATG 526

RESULT 12
US-08-318-970B-38
Sequence 38, Application US/08318970B
Patent No. 3589573
GENERAL INFORMATION:
APPLICANT: Hideaki HAGIWARA, et al.
TITLE OF INVENTION: AMINO ACID SEQUENCES OF ANTI-IDIOTYPIC
TITLE OF INVENTION: ANTIBODIES AGAINST ANTI-CANCER HUMAN MONOCLONAL ANTIBODY
TITLE OF INVENTION: AND DNA BASE SEQUENCES ENCODING THOSE SEQUENCES
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Office of Sherman and Shalloway
STREET: 413 N. Washington Street
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: Dell System 210; Intel 80 285 Microprocessor
OPERATING SYSTEM: MS DOS 3.3; Version 5.1
SOFTWARE: Word Perfect, Version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,970B
FILING DATE: October 6, 1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Richard A. Steinberg
REGISTRATION NUMBER: 26,588
REFERENCE/DOCKET NUMBER: S-2371
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 549-2282
TELEFAX: (703) 836-0106
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: mRNA
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: mouse
FEATURE:
NAME/KEY: Idio 27 L chain
US-08-318-970B-38

Query Match 1.4%; Score 17; DB 1; Length 321;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 794 TCACCATCAGCGGTGTG 810
Db 218 TCACCATCAGCGGTGTG 234

RESULT 13
US-08-318-970B-47
Sequence 47, Application US/08318970B
Patent No. 5589573
GENERAL INFORMATION:
APPLICANT: Hideaki HAGIWARA, et al.
TITLE OF INVENTION: AMINO ACID SEQUENCES OF ANTI-IDIOTYPIC
TITLE OF INVENTION: ANTIBODIES AGAINST ANTI-CANCER HUMAN MONOCLONAL ANTIBODY
TITLE OF INVENTION: AND DNA BASE SEQUENCES ENCODING THOSE SEQUENCES
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Office of Sherman and Shalloway
STREET: 413 N. Washington Street
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: Dell System 210; Intel 80 285 Microprocessor
OPERATING SYSTEM: MS DOS 3.3
SOFTWARE: Word Perfect, Version 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,970B
FILING DATE: October 6, 1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Richard A. Steinberg
REGISTRATION NUMBER: 26,588
REFERENCE/DOCKET NUMBER: S-2371
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 549-2282
TELEFAX: (703) 836-0106
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 420
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: mouse
FEATURE:
NAME/KEY: Clone 27KA2
US-08-318-970B-47

Query Match 1.4%; Score 17; DB 1; Length 420;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 794 TCACCATCAGCGGTGTG 810
DB 308 TCACCATCAGCGGTGTG 324

RESULT 14
US-09-359-301A-24
; Sequence 24, Application US/09359301A
; Patent No. 6426185
; GENERAL INFORMATION:
; APPLICANT: Kumagai, Monto H.
; APPLICANT: della-Cioppa, Guy R.
; APPLICANT: Erwin, Robert L.
; APPLICANT: McGee, David R.
; TITLE OF INVENTION: METHOD OF DETERMINING THE PRESENCE OF A
; TITLE OF INVENTION: TRAIT IN A PLANT BY TRANSFECTING A NUCLEIC ACID SEQUENCE OF
; TITLE OF INVENTION: A DONOR PLANT INTO A DIFFERENT HOST PLANT IN AN ANTI-SENSE
; FILE REFERENCE: 008010137US04
; CURRENT APPLICATION NUMBER: US/09/359,301A
; CURRENT FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-359-301A-24

Query Match 1.4%; Score 17; DB 4; Length 550;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 ACCGTGCTGAGATCACC 87
DB 406 ACCGTGCTGAGATCACC 422

RESULT 15
US-09-071-035-13/c
; Sequence 13, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 555 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-071-035-13

Query Match 1.4%; Score 17; DB 4; Length 555;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 343 AAGAGCAATGCTGCAA 359
DB 60 AAGAGCAATGCTGCAA 44

Search completed: February 5, 2004, 08:14:30
Job time: 97 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2004, 06:01:29 ; Search time 489 Seconds
(without alignments)
8836.199 Million cell updates/sec

Title: US-09-643-755B-1

Perfect score: 1173
Sequence: 1 atgacttccttaagtcctt.....ggctacgaacgaatcga 1173

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 2449703 seqs, 1841816367 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	9.0	1041	US-10-324-131-10	Sequence 10, Appl
2	49	4.2	2733	US-09-893-525-6	Sequence 6, Appl
3	34	2.9	1096	US-10-322-746-1	Sequence 1, Appl
4	33	2.8	1119	US-10-322-746-3	Sequence 3, Appl
5	23	2.0	609	US-09-770-149-886	Sequence 886, App
6	23	2.0	1521	US-09-938-842A-2245	Sequence 2245, Ap
7	23	2.0	1053	US-09-938-842A-2245	Sequence 2245, Ap
8	22	1.9	1080	US-10-322-746-16	Sequence 16, Appl
9	22	1.9	1101	US-10-322-746-14	Sequence 14, Appl
10	22	1.9	1134	US-10-322-746-12	Sequence 12, Appl
11	22	1.9	2756	US-09-887-576-449	Sequence 449, App
12	19	1.6	231	US-09-815-343-50	Sequence 50, Appl
13	19	1.6	503	US-10-027-632-106390	Sequence 106390,
14	19	1.6	503	US-10-027-632-106390	Sequence 106390,
15	19	1.6	503	US-10-027-632-106390	Sequence 106390,

16	19	1.6	556	10	US-09-998-598-1222	Sequence 1222, Ap
17	19	1.6	689	11	US-09-809-391-263	Sequence 263, App
18	19	1.6	689	13	US-09-882-171-263	Sequence 263, App
19	19	1.6	714	13	US-10-027-632-22346	Sequence 22346, A
20	19	1.6	714	14	US-10-027-632-22346	Sequence 22346, A
21	19	1.6	1751	11	US-09-809-391-110	Sequence 110, App
22	19	1.6	1751	13	US-09-882-171-110	Sequence 110, App
23	19	1.6	4170	11	US-09-919-039-221	Sequence 221, App
24	19	1.6	4170	15	US-10-168-425-23	Sequence 23, Appl
25	19	1.6	13440	15	US-10-213-948-12	Sequence 12, Appl
26	19	1.6	13953	14	US-10-096-861-3	Sequence 3, Appl
27	19	1.5	60	13	US-09-908-975-23544	Sequence 23544, A
28	18	1.5	527	11	US-09-918-995-31599	Sequence 31599, A
29	18	1.5	558	13	US-10-027-632-207782	Sequence 207782,
30	18	1.5	558	13	US-10-027-632-207782	Sequence 207784,
31	18	1.5	558	13	US-10-027-632-207785	Sequence 207785,
32	18	1.5	558	14	US-10-027-632-207782	Sequence 207782,
33	18	1.5	558	14	US-10-027-632-207784	Sequence 207784,
34	18	1.5	558	14	US-10-027-632-207785	Sequence 207785,
35	18	1.5	632	13	US-10-027-632-200742	Sequence 200742,
36	18	1.5	632	13	US-10-027-632-200742	Sequence 200743,
37	18	1.5	632	14	US-10-027-632-200742	Sequence 200743,
38	18	1.5	632	14	US-10-027-632-200743	Sequence 200743,
39	18	1.5	648	11	US-09-791-279-83	Sequence 83, Appl
40	18	1.5	1117	13	US-10-027-632-207783	Sequence 207783,
41	18	1.5	1117	14	US-10-027-632-207783	Sequence 207783,
42	18	1.5	1365	13	US-10-027-632-251427	Sequence 251427,
43	18	1.5	1365	13	US-10-027-632-251428	Sequence 251428,
44	18	1.5	1365	13	US-10-027-632-251429	Sequence 251429,
45	18	1.5	1365	13	US-10-027-632-251430	Sequence 251430,

ALIGNMENTS

RESULT 1
US-10-324-131-10
Sequence 10, Application US/10324131
Publication No. US2003017537A1
GENERAL INFORMATION:
APPLICANT: Moloney, Maurice M.
TITLE OF INVENTION: Expression of Epidermal Growth Factor in Plant Seeds
FILE REFERENCE: 9369-251
CURRENT APPLICATION NUMBER: US/10/324,131
PRIOR FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US 07/659,835
PRIOR FILING DATE: 1991-02-22
PRIOR APPLICATION NUMBER: US 08/142,418
PRIOR FILING DATE: 1993-11-16
PRIOR APPLICATION NUMBER: US 08/366,783
PRIOR FILING DATE: 1994-12-30
PRIOR APPLICATION NUMBER: US 08/846,021
PRIOR FILING DATE: 1997-04-25
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 1041
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oleosin-BGF
US-10-324-131-10

Query Match 9.0%; Score 105; DB 13; Length 1041;
Best Local Similarity 100.0%; Pred. No. 4.7e-47;
Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 GCTGAGATCACCAGCCGATCTCTCTACCAAGTAAGTCTTCCTGTAAGCCGCTGAAGAA 135
DB 763 GCTGAGATCACCAGCCGATCTCTCTACCAAGTAAGTCTTCCTGTAAGCCGCTGAAGAA 822
QY 136 CATGACTTCTAGAAAGCTCTTCGAGAAACAAGTATGATC 180

Db 823 CATGACCTTACAGACGCTTCTTCGAGAAACACAGATGACATC 867

RESULT 2

US-09-893-525-6
; Sequence 6, Application US/09893525
; Publication No. US20030126631A1
; GENERAL INFORMATION:
; APPLICANT: Moloney, Maurice M.
; APPLICANT: van Rooijen, Gijb
; TITLE OF INVENTION: Preparation of Heterologous Proteins on Oil Bodies
; FILE REFERENCE: 9369-172
; CURRENT APPLICATION NUMBER: US/09/893,525
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 09/210,843
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: US 08/846,021
; PRIOR FILING DATE: 1997-04-25
; PRIOR APPLICATION NUMBER: US 08/366,783
; PRIOR FILING DATE: 1994-12-30
; PRIOR APPLICATION NUMBER: US 08/142,418
; PRIOR FILING DATE: 1993-11-16
; PRIOR APPLICATION NUMBER: US 07/659,835
; PRIOR FILING DATE: 1991-02-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 2733
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (850)..(1203)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (1444)..(2727)
; OTHER INFORMATION:
; US-09-893-525-6

Query Match 4.2%; Score 49; DB 11; Length 2733;
Best Local Similarity 100.0%; Pred. No. 3.9e-16;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 656 TGTTCGCTTACATGACAGAGATGCCAGAGACATGCTCAGCGCT 704
Db 2210 TGTTCGCTTACATGACAGAGATGCCAGAGACATGCTCAGCGCT 2258

RESULT 3

US-10-322-746-1
; Sequence 1, Application US/10322746
; Publication No. US20030166162A1
; GENERAL INFORMATION:
; APPLICANT: van Rooijen, Gijb
; APPLICANT: Alcantara, Joemel
; APPLICANT: Moloney, Maurice M.
; TITLE OF INVENTION: Method for Cleavage of Fusion Proteins
; FILE REFERENCE: 9369-227
; CURRENT APPLICATION NUMBER: US/10/322,746
; CURRENT FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1096
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GST-Pro-Hirudin fusion
; US-10-322-746-1

Query Match 2.9%; Score 34; DB 13; Length 1096;

Best Local Similarity 100.0%; Pred. No. 7.6e-08;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 CAGTATGCATCAGCAGCAAGTACTCCGGCTTCG 202
Db 802 CAGTATGCATCAGCAGCAAGTACTCCGGCTTCG 835

RESULT 4

US-10-322-746-3
; Sequence 3, Application US/10322746
; Publication No. US20030166162A1
; GENERAL INFORMATION:
; APPLICANT: van Rooijen, Gijb
; APPLICANT: Alcantara, Joemel
; APPLICANT: Moloney, Maurice M.
; TITLE OF INVENTION: Method for Cleavage of Fusion Proteins
; FILE REFERENCE: 9369-227
; CURRENT APPLICATION NUMBER: US/10/322,746
; CURRENT FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: His-Pro-cGH
; US-10-322-746-3

Query Match 2.8%; Score 33; DB 13; Length 1119;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 169 CAGTATGCATCAGCAGCAAGTACTCCGGCTTCG 201
Db 220 CAGTATGCATCAGCAGCAAGTACTCCGGCTTCG 252

RESULT 5

US-09-770-149-886/C
; Sequence 886, Application US/09770149
; Patent No. US2002059663A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jörn
; APPLICANT: An, Yong-Olang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2024 (PAPA-013PRV)
; CURRENT APPLICATION NUMBER: US/09/770,149
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,506
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 886

LENGTH: 609
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-149-886

Query Match 2.0%; Score 23; DB 9; Length 609;
Best Local Similarity 100.0%; Pred. No. 0.091;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 TGGATCTTGGGAGATGTTCAT 1109
DB 389 TGGATCTTGGGAGATGTTCAT 367

RESULT 6
US-09-938-842A-2245

Sequence 2245, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2245
LENGTH: 1521
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2245

Query Match 2.0%; Score 23; DB 10; Length 1521;
Best Local Similarity 100.0%; Pred. No. 0.091;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 TGGATCTTGGGAGATGTTCAT 1109
DB 1435 TGGATCTTGGGAGATGTTCAT 1457

RESULT 7
US-09-938-842A-2245

Sequence 2245, Application US/09938842A
Publication No. US20040009476A9
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2245
LENGTH: 1521

TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2245

Query Match 2.0%; Score 23; DB 12; Length 1521;
Best Local Similarity 100.0%; Pred. No. 0.091;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 TGGATCTTGGGAGATGTTCAT 1109
DB 1435 TGGATCTTGGGAGATGTTCAT 1457

RESULT 8
US-10-322-746-18

Sequence 18, Application US/10322746
Publication No. US20030166162A1
GENERAL INFORMATION:
APPLICANT: van Rooijen, Gijb
APPLICANT: Alcantara, Joanel
APPLICANT: Moloney, Maurice M.
TITLE OF INVENTION: Method for Cleavage of Fusion Proteins
FILE REFERENCE: 9369-227
CURRENT APPLICATION NUMBER: US/10/322,746
CURRENT FILING DATE: 2002-12-19
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.1
SEQ ID NO 18
LENGTH: 1053
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: GST-KLIP14-cystatin
US-10-322-746-18

Query Match 1.9%; Score 22; DB 13; Length 1053;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 CAGTATGGCATCAGCAGCACT 190
DB 721 CAGTATGGCATCAGCAGCACT 742

RESULT 9
US-10-322-746-16

Sequence 16, Application US/10322746
Publication No. US20030166162A1
GENERAL INFORMATION:
APPLICANT: van Rooijen, Gijb
APPLICANT: Alcantara, Joanel
APPLICANT: Moloney, Maurice M.
TITLE OF INVENTION: Method for Cleavage of Fusion Proteins
FILE REFERENCE: 9369-227
CURRENT APPLICATION NUMBER: US/10/322,746
CURRENT FILING DATE: 2002-12-19
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.1
SEQ ID NO 16
LENGTH: 1080
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: GST-KLIP12-cystatin
US-10-322-746-16

Query Match 1.9%; Score 22; DB 13; Length 1080;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 CAGTATGGCATCAGCAGCACT 190
DB 748 CAGTATGGCATCAGCAGCACT 769

RESULT 10
US-10-322-746-14
; Sequence 14, Application US/10322746
; Publication No. US20030166162A1
; GENERAL INFORMATION:
; APPLICANT: van Rooijen, Gijb
; APPLICANT: Alcantara, Joemel
; APPLICANT: Moloney, Maurice M.
; TITLE OF INVENTION: Method for Cleavage of Fusion Proteins
; FILE REFERENCE: 9369-227
; CURRENT APPLICATION NUMBER: US/10/322,746
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 14
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GST-KLIP11-cystatin
US-10-322-746-14

Query Match 1.9%; Score 22; DB 13; Length 1101;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 CAGTATGCATCAGCAGCAAGT 190
Db 769 CAGTATGCATCAGCAGCAAGT 790

RESULT 11
US-10-322-746-12
; Sequence 12, Application US/10322746
; Publication No. US20030166162A1
; GENERAL INFORMATION:
; APPLICANT: van Rooijen, Gijb
; APPLICANT: Alcantara, Joemel
; APPLICANT: Moloney, Maurice M.
; TITLE OF INVENTION: Method for Cleavage of Fusion Proteins
; FILE REFERENCE: 9369-227
; CURRENT APPLICATION NUMBER: US/10/322,746
; CURRENT FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 12
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GST-KLIP4-cystatin
US-10-322-746-12

Query Match 1.9%; Score 22; DB 13; Length 1134;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 CAGTATGCATCAGCAGCAAGT 190
Db 802 CAGTATGCATCAGCAGCAAGT 823

RESULT 12
US-09-887-576-449
; Sequence 449, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.

; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 449
; LENGTH: 2756
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-449

Query Match 1.9%; Score 22; DB 10; Length 2756;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1088 GGATCTTGGGAGATGTGTTTCAT 1109
Db 2671 GGATCTTGGGAGATGTGTTTCAT 2692

RESULT 13
US-09-815-343-50/c
; Sequence 50, Application US/09815343
; Patent No. US2001005596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jianshun
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815,343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(231)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-343-50

Query Match 1.6%; Score 19; DB 9; Length 231;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1072 AACCATTCGAGAAATGGA 1090
Db 153 AACCATTCGAGAAATGGA 135

RESULT 14
US-10-027-632-106390
; Sequence 106390, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129

Search completed: February 5, 2004, 08:25:27
Job time : 494 secs

CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 106390
LENGTH: 503
TYPE: DNA
ORGANISM: Human
US-10-027-632-106390

Query Match 1.6%; Score 19; DB 13; Length 503;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 GCATTCCTCTCTACAAAG 107
|||||
Db 436 GCATTCCTCTCTACAAAG 454

RESULT 15
US-10-027-632-106390
Sequence 106390, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 106390
LENGTH: 503
TYPE: DNA
ORGANISM: Human
US-10-027-632-106390

Query Match 1.6%; Score 19; DB 14; Length 503;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 GCATTCCTCTCTACAAAG 107
|||||
Db 436 GCATTCCTCTCTACAAAG 454

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2004, 05:19:09 ; Search time 2885 Seconds
(without alignments)
9881.855 Million cell updates/sec

Title: US-09-643-755B-1

Perfect score: 1173
Sequence: 1 atgaactccttaagctctt.....ggtcagcctaagaacatctga 1173

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrt1:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	4.2	307	14	CB221243
2	49	4.2	383	12	BG938086
3	49	4.2	430	12	BG937697
4	49	4.2	472	12	BG938320

5	43	3.7	399	12	BG937723
6	42	3.6	410	14	CB220875
7	38	3.2	297	14	CB220520
8	38	3.2	443	14	CB220074
9	32	2.7	204	29	BZ911661
10	23	2.0	325	9	AV532492
11	23	2.0	339	14	T20903
12	23	2.0	380	9	AV538262
13	23	2.0	391	9	AA969042
14	23	2.0	396	9	AV801400
15	23	2.0	398	9	AV801901
16	23	2.0	398	9	AV802605
17	23	2.0	404	9	AA411567
18	23	2.0	405	9	AV806186
19	23	2.0	405	9	AV814825
20	23	2.0	408	9	AV518232
21	23	2.0	408	9	AV798981
22	23	2.0	409	9	AV806470
23	23	2.0	410	9	AV802625
24	23	2.0	410	9	AV803851
25	23	2.0	411	9	AV559941
26	23	2.0	413	9	AV802692
27	23	2.0	420	9	AV805687
28	23	2.0	421	9	AV788245
29	23	2.0	421	9	AV788317
30	23	2.0	424	9	AV796683
31	23	2.0	424	9	AV820962
32	23	2.0	425	9	AV796827
33	23	2.0	430	9	AV810526
34	23	2.0	431	9	AV810327
35	23	2.0	433	9	AV798761
36	23	2.0	434	9	AV797571
37	23	2.0	438	9	AV802780
38	23	2.0	438	9	AA411566
39	23	2.0	441	9	AV792423
40	23	2.0	442	13	BX096065
41	23	2.0	443	9	AV815805
42	23	2.0	447	9	AV793780
43	23	2.0	455	9	AV794649
44	23	2.0	483	9	AV536021
45	23	2.0	505	9	AI479358

ALIGNMENTS

RESULT 1
LOCUS CB221243
DEFINITION IABO32A10 Bos taurus Abomasum #1 library Bos taurus CDNA, mRNA
sequence.
ACCESSION CB221243
VERSION CB221243.1 GI:28291757
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos.
REFERENCE 1 (bases 1 to 307)
Hansen, C., Fu, A., Meng, Y., Li, C., Okine, E., Sensen, C.W., Gordon, P.M.K. and Moore, S.S.
Gene Expression Profiling of the Bovine Gastrointestinal Tract
Unpublished
Contact: Dr. Stephen Moore
Beef Genomics Laboratory
Dept of AFNS, University of Alberta
410 Agri/For. Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
Tel: 780 492 0169
Fax: 780 492 4265
Email: stephen.moore@ualberta.ca
Insert Length: 307 Std Error: 0.00
POLYA=No.

FEATURES
source
1.307
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="Smooth muscle"
/cell_type="Epithelial"
/dev_stage="Young adult"
/lab_host="XLI-BlueMRF"-strain"
/clone_lib="Bos taurus Abomasum #1 library"
/note="Organ: Abomasum; Vector: Uni-2ZAPXR; Site_1: Ecor
I; Site_2: Xho I"
62 a 91 c 93 g 61 t

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 49; DB 14; Length 307;
Pred. No. 4.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 656 TGTTCGCGTTTACATGACAGGAATGCCAGAGACATGCTCAGCCT 704
Db 65 TGTTCGCGTTTACATGACAGGAATGCCAGAGACATGCTCAGCCT 113

RESULT 2
BG938086 383 bp mRNA linear EST 11-JUN-2001
LOCUS 1AB011A08 Bovine Abomasum cDNA library Bos taurus cDNA 5', mRNA
DEFINITION
ACCESSION BG938086
VERSION BG938086.1 GI:14337458
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 383)
Email: stephen.moore@ualberta.ca
The sequence best matches gb:BP019786 (Bos primigenius prochimysin
mRNA, complete cds) in main database at high score of 735.0 and
E-value of 0.0
PCR Primers
FORWARD: M13 Forward
BACKWARD: M13 Reverse
Seq primer: T3 primer
High quality sequence stop: 383
POLYA=No.

FEATURES
source
1.383
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/sex="Two males and one female mixed"
/tissue_type="Gastrointestinal tissue (GIT)"
/cell_type="Epithelial"
/dev_stage="Young adult"
/lab_host="XLI-BlueMRF"-strain"
/clone_lib="Bovine Abomasum cDNA library"
/note="Organ: Abomasum; Vector: Uni-2ZAPXR; Site_1: Ecor
I; Site_2: Xho I"
83 a 111 c 113 g 76 t

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 49; DB 12; Length 383;
Pred. No. 5.2e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 656 TGTTCGCGTTTACATGACAGGAATGCCAGAGACATGCTCAGCCT 704
Db 58 TGTTCGCGTTTACATGACAGGAATGCCAGAGACATGCTCAGCCT 106

RESULT 3
BG937697 430 bp mRNA linear EST 11-JUN-2001
LOCUS 1AB005E01 Bovine Abomasum cDNA library Bos taurus cDNA 5', mRNA
DEFINITION
ACCESSION BG937697
VERSION BG937697.1 GI:14337069
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 430)
Email: stephen.moore@ualberta.ca
The sequence best matches gb:BOVCHYMOA (bovine chymosin a (rennin)
mRNA) in main database at high score of 844.0 and E-value of 0.0
PCR Primers
FORWARD: M13 Forward
BACKWARD: M13 Reverse
Seq primer: T3 primer
High quality sequence stop: 430
POLYA=No.

FEATURES
source
1.430
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/sex="Two males and one female mixed"
/tissue_type="Gastrointestinal tissue (GIT)"
/cell_type="Epithelial"
/dev_stage="Young adult"
/lab_host="XLI-BlueMRF"-strain"
/clone_lib="Bovine Abomasum cDNA library"
/note="Organ: Abomasum; Vector: Uni-2ZAPXR; Site_1: Ecor
I; Site_2: Xho I"
97 a 130 c 119 g 84 t

BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 49; DB 12; Length 430;
Pred. No. 5.4e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 656 TGTTCGCGTTTACATGACAGGAATGCCAGAGACATGCTCAGCCT 704
Db 52 TGTTCGCGTTTACATGACAGGAATGCCAGAGACATGCTCAGCCT 100

RESULT 4
BG938320 472 bp mRNA linear EST 11-JUN-2001
LOCUS 1AB015E12 Bovine Abomasum cDNA library Bos taurus cDNA 5', mRNA
DEFINITION
ACCESSION BG938320
VERSION BG938320.1 GI:14337692

KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 472)
AUTHORS Moore, S.S., Hansen, C., Li, C., Fu, A., Meng, Y. and Li, G.
TITLE cDNA's from bovine abomasum tissue
JOURNAL Unpublished
COMMENT Contact: Dr. Stephen Moore
Beef Genomics Laboratory
Dept of AFNS, University of Alberta
410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
Tel: 780 492 0169
Fax: 780 492 4265
Email: stephen.moore@ualberta.ca
The sequence best matches gb:BOVCHYMOA (bovine chymosin a (remnin
mrna) in main database at high score of 928.0 and E-value of 0.0
PCR primers
FORWARD: M13 Forward
BACKWARD: M13 Reverse
Seq primer: T3 primer
High quality sequence stop: 472
POLYA=No.

FEATURES
source Location/Qualifiers
1..472
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/sex="Two males and one female mixed"
/tissue_type="Gastrointestinal tissue (GIT)"
/cell_type="Epithelial"
/dev_stage="Young adult"
/lab_host="XLI-BlueMRF"-strain"
/clone_lib="Bovine Abomasum cDNA library"
/note="Organ: Abomasum; Vector: Uni-22APXR; Site_1: Ecor
I; Site_2: Xho I"
BASE COUNT 104 a 143 c 131 g 94 t
ORIGIN

Query Match 4.2%; Score 49; DB 12; Length 472;
Best Local Similarity 100.0%; Pred. No. 5.6e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 656 TGTTCGCTTACATGACAGCAATGGCCAGAGACATGCTCAGCT 704
|||||
65 TGTTCGCTTACATGACAGCAATGGCCAGAGACATGCTCAGCT 113
|||||

RESULT 5
BG937723 399 bp mRNA linear EST 11-JUN-2001
LOCUS 1ABO05D06 Bovine Abomasum cDNA Library Bos taurus cDNA 5', mRNA
DEFINITION
ACCESSION BG937723
VERSION BG937723
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 399)
AUTHORS Moore, S.S., Hansen, C., Li, C., Fu, A., Meng, Y. and Li, G.
TITLE cDNA's from bovine abomasum tissue
JOURNAL Unpublished
COMMENT Contact: Dr. Stephen Moore
Beef Genomics Laboratory
Dept of AFNS, University of Alberta
410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
Tel: 780 492 0169
Fax: 780 492 4265

Email: stephen.moore@ualberta.ca
The sequence best matches gb:BPUL9786 (Bos primigenius prochymosin
mRNA, complete cds) in main database at high score of 767.0 and
E-value of 0.0
PCR primers
FORWARD: M13 Forward
BACKWARD: M13 Reverse
Seq primer: T3 primer
High quality sequence stop: 399
POLYA=No.

FEATURES
source Location/Qualifiers
1..399
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/sex="Two males and one female mixed"
/tissue_type="Gastrointestinal tissue (GIT)"
/cell_type="Epithelial"
/dev_stage="Young adult"
/lab_host="XLI-BlueMRF"-strain"
/clone_lib="Bovine Abomasum cDNA library"
/note="Organ: Abomasum; Vector: Uni-22APXR; Site_1: Ecor
I; Site_2: Xho I"
BASE COUNT 89 a 121 c 112 g 77 t
ORIGIN

Query Match 3.7%; Score 43; DB 12; Length 399;
Best Local Similarity 100.0%; Pred. No. 7e-11;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 662 CGGTTACATGACAGCAATGCCAGAGACATGCTCAGCT 704
|||||
1 CGGTTACATGACAGCAATGCCAGAGACATGCTCAGCT 43
|||||

RESULT 6
CB220875 410 bp mRNA linear EST 10-FEB-2003
LOCUS 1ABO27D10 Bos taurus Abomasum #1 library Bos taurus cDNA, mRNA
DEFINITION
ACCESSION CB220875
VERSION CB220875.1 GI:28291389
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 410)
AUTHORS Hansen, C., Fu, A., Meng, Y., Li, C., Okine, E., Senses, C.W., Gordon
P.M.K. and Moore, S.S.
TITLE Gene Expression Profiling of the Bovine Gastrointestinal Tract
JOURNAL Unpublished
COMMENT Contact: Dr. Stephen Moore
Beef Genomics Laboratory
Dept of AFNS, University of Alberta
410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
Tel: 780 492 0169
Fax: 780 492 4265
Email: stephen.moore@ualberta.ca
Insert length: 410 Std Error: 0.00
POLYA=No.

FEATURES
source Location/Qualifiers
1..410
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="Smooth muscle"
/cell_type="Epithelial"
/dev_stage="Young adult"
/lab_host="XLI-BlueMRF"-strain"
/clone_lib="Bos taurus Abomasum #1 library"
/note="Organ: Abomasum; Vector: Uni-22APXR; Site_1: Ecor
I; Site_2: Xho I"

BASE COUNT 90 a 123 c 120 g 77 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 2.4e-10;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 663 GGTTCATGACGACGAGATGGCCAGAGACATGCTCAGCCT 704
Db 36 GGTTCATGACGACGAGATGGCCAGAGACATGCTCAGCCT 77

RESULT 7
CB220520 297 bp mRNA linear EST 10-FEB-2003
LOCUS 1Ab0231a12 Bos taurus Abomasum #1 library Bos taurus cDNA, mRNA
DEFINITION
sequence.
ACCESSION CB220520
VERSION CB220520.1 GI:28291034
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 297)
Hansen, C., Fu, A., Meng, Y., Li, C., Okine, E., Sensen, C.W., Gordon
P.M.K. and Moore, S.S.
Gene Expression Profiling of the Bovine Gastrointestinal Tract
Unpublished
Contact: Dr. Stephen Moore
Beef Genomics Laboratory
Dept of AFNS, University of Alberta
410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
Tel: 780 492 0169
Fax: 780 492 4265
Email: stephen.moore@ualberta.ca
Insert Length: 297 Std Error: 0.00
POLYA=No.

FEATURES
source Location/Qualifiers
1..297
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="Smooth muscle"
/cell_type="Epithelial"
/dev_stage="Young adult"
/lab_host="X11-BlueMRP-strain"
/clone_1lb="Bos taurus Abomasum #1 library"
/note="Organ: Abomasum; Vector: Uni-2ZAPXR; site_1: Ecor
I; site_2: Xho I"

BASE COUNT 67 a 91 c 81 g 58 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 778 TTCACGTGACGACGTTCACCATCAGCGGTGTGTTCT 815
Db 16 TTCACGTGACGACGTTCACCATCAGCGGTGTGTTCT 53

RESULT 8
CB220074 443 bp mRNA linear EST 10-FEB-2003
LOCUS 1Ab017B12 Bos taurus Abomasum #1 library Bos taurus cDNA, mRNA
DEFINITION
sequence.
ACCESSION CB220074
VERSION CB220074.1 GI:28290568
KEYWORDS EST.
SOURCE Bos taurus (cow)

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 443)
Hansen, C., Fu, A., Meng, Y., Li, C., Okine, E., Sensen, C.W., Gordon
P.M.K. and Moore, S.S.
Gene Expression Profiling of the Bovine Gastrointestinal Tract
Unpublished
Contact: Dr. Stephen Moore
Beef Genomics Laboratory
Dept of AFNS, University of Alberta
410 Agri/For, Dept of AFNS, U of A, Edmonton, AB, T6G 2P5, Canada
Tel: 780 492 0169
Fax: 780 492 4265
Email: stephen.moore@ualberta.ca
Insert Length: 443 Std Error: 0.00
POLYA=Yes.

FEATURES
source Location/Qualifiers
1..443
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="Smooth muscle"
/cell_type="Epithelial"
/dev_stage="Young adult"
/lab_host="X11-BlueMRP-strain"
/clone_1lb="Bos taurus Abomasum #1 library"
/note="Organ: Abomasum; Vector: Uni-2ZAPXR; site_1: Ecor
I; site_2: Xho I"

BASE COUNT 125 a 132 c 96 g 90 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 907 ATTGAGCCACACAGACCACTACGCTGAGTTGACAT 944
Db 37 ATTGAGCCACACAGAACCACTACGCTGAGTTGACAT 74

RESULT 9
BZ911661 204 bp DNA linear GSS 12-JUN-2003
LOCUS CH240_109D12.TV CHOR1-240 Bos taurus genomic clone CH240_109D12,
DEFINITION genomic survey sequence.
ACCESSION BZ911661
VERSION BZ911661.1 GI:31637047
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 204)
Larkin, D.M., Everts-van der Wind, A., Rebeiz, M., Schweitzer, P.,
Bachman, S., Green, S., Campos, E.J., Benson, L.D., Edwards, J., Liu, L.,
Mowack, J.E., de Jong, P.J. and Lewin, H.A.
A Cattle-Human Comparative Map Built with Cattle BAC-ends and Human
Genome Sequence
Unpublished
Other GSSs: CH240_109D12.TV
Contact: Harris Lewin
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Clones are derived from the bovine BAC library CHOR1-240
(http://www.cho1.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/ordering/information.htm>). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBMC) by the University of Illinois at Urbana-Champaign, USA with funds provided by grant No. AG202-34480-11828 from USDA-CSRES and AG99-35205-8534 from USDA/NRI (Livestock Genome Sequencing Initiative)
 Plate: 109 row: D column: 12
 Seq primer: SP6
 Class: BAC ends.

FEATURES

Source

Location/Qualifiers
 1. .204
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="bred: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_109D12"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: Mbol; Site 2: Mbol; Hereford bull Li Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"
 BASE COUNT 42 a 60 c 54 g 43 t
 ORIGIN

Query Match 2.7%; Score 32; DB 29; Length 204;
 Best Local Similarity 100.0%; Pred. No. 3e-05;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 169 CAGTATGGCATCAGCAGCACTACCGGCTT 200
 |||||
 22 CAGTATGGCATCAGCAGCACTACCGGCTT 53

RESULT 10
 AV532492/c 325 bp mRNA linear EST 01-SEP-2000
 LOCUS AV532492 Arabidopsis thaliana flower buds Columbia Arabidopsis
 DEFINITION thaliana cDNA clone FB043b03f 3', mRNA sequence.
 ACCESSION AV532492
 VERSION AV532492.1 GI:8692775
 KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE
 AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
 TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
 JOURNAL 20363093
 PUBMED 10907847

COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES
 source Location/Qualifiers
 1. .325
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="FB043b03f"
 /tissue_type="flower buds"
 /clone_lib="Arabidopsis thaliana flower buds Columbia"
 /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2: XhoI"

BASE COUNT 106 a 92 c 59 g 68 t

ORIGIN

Query Match 2.0%; Score 23; DB 9; Length 325;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 TGGATCTTGGAGATGTGTTTCAT 1109
 |||||
 DB 209 TGGATCTTGGAGATGTGTTTCAT 187

RESULT 11

T20903 339 bp mRNA linear EST 07-JAN-1998
 LOCUS T20903
 DEFINITION 2911 lambda-PRL2 Arabidopsis thaliana cDNA clone 89P9T7, mRNA sequence.

ACCESSION T20903
 VERSION T20903.1 GI:2756823
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE
 AUTHORS Newman, T., deBruin, J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E., and Somerville, C.
 TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 JOURNAL Plant Physiol. 106, 1241-1255 (1994)
 MEDLINE 95148729
 PUBMED 7846151

COMMENT On Jan 7, 1998 this sequence version replaced gi:502344.
 Contact: Thomas Newman
 MSU-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313lc@lhm.cl.msu.edu
 Seg primer: T7.

FEATURES
 source Location/Qualifiers
 1. .339
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone="89P9T7"
 /clone_lib="lambda-PRL2"
 /note="Vector: lambda Z1p-Lox; Site 1: Sal; Site 2: Not; lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Z1p-Lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dT primed cDNA."

BASE COUNT 67 a 61 c 92 g 106 t 13 others
 ORIGIN

Query Match 2.0%; Score 23; DB 14; Length 339;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 TGGATCTTGGAGATGTGTTTCAT 1109
 |||||
 DB 82 TGGATCTTGGAGATGTGTTTCAT 104

```

RESULT 12
AVS38262/c 380 bp mRNA linear EST 06-SEP-2000
LOCUS AVS38262 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
DEFINITION AVS38262 cDNA clone R2113a03f 3', mRNA sequence.
ACCESSION AVS38262
VERSION AVS38262.1 GI:8698545
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 380)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20365093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizukazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
FEATURES
source
1..380
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="R2113a03f"
/tissue_type="roots"
/clone_1ib="Arabidopsis thaliana roots Columbia"
/note="Vector: pBluescriptII SK-, Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 130 a 96 c 65 g 89 t
ORIGIN
Query Match 2.0%; Score 23; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1087 TGGATCTTGGAGATGTTTCAT 1109
|||||
Db 332 TGGATCTTGGAGATGTTTCAT 310
|||||

RESULT 13
AA969042/c 391 bp mRNA linear EST 07-JUL-1998
LOCUS AA969042 GBC S1 Homo sapiens cDNA clone
DEFINITION IMAGE:1579597 3' similar to TR:Q27951 Q27951 PROCHYMOSIN. ; mRNA
SEQUENCE.
ACCESSION AA969042
VERSION AA969042.1 GI:3144222
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 391)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-rc@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 619 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham

```

```

FEATURES
source
High quality sequence stop: 378.
Location/Qualifiers
1..391
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1579597"
/lab_host="DH10B"
/clone_1ib="Soares_NFL_T_GBC_S1"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CCAP GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 725096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 82 a 101 c 115 g 93 t
ORIGIN
Query Match 2.0%; Score 23; DB 9; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1015 CCCTCCGCTATACGACGAGA 1037
|||||
Db 256 CCCTCCGCTATACGACGAGA 234
|||||

RESULT 14
AV801400/c 396 bp mRNA linear EST 29-MAR-2002
LOCUS AV801400 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-27-P07 3',
DEFINITION AV801400 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-27-P07 3',
mRNA sequence.
ACCESSION AV801400
VERSION AV801400.1 GI:19835385
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 396)
AUTHORS Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Saito, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, T., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.,
and Shinozaki, K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@rc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified lambda PhC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.
FEATURES
source
Location/Qualifiers
1..396
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL09-27-P07"

```

/dev stage="plants at various developmental stages from
germination to mature seeds"
/lab host="DH10B"
/clone lib="RAFL9"
/note="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

BASE COUNT 133 a 101 c 70 g 92 t

ORIGIN

Query Match 2.0%; Score 23; DB 9; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1087 TGGATCTTGGGAGATGTTTCAT 1109
|||||
Db 330 TGGATCTTGGGAGATGTTTCAT 308

RESULT 15
AV801901/c 398 bp mRNA linear EST 29-MAR-2002
DEFINITION AV801901 RAF19 Arabidopsis thaliana cDNA clone RAF109-29-P20 3',
mRNA sequence.
ACCESSION AV801901
VERSION AV801901.1 GI:19835886
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
1 (bases 1 to 398)
; euroside II; Brassicales; Brassicaceae; Arabidopsis.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished

TITLE
JOURNAL
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@tc.riken.go.jp

FEATURES
source 1..398
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL09-29-P20"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab host="DH10B"
/clone lib="RAFL9"
/note="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

BASE COUNT 133 a 103 c 71 g 91 t

ORIGIN

Query Match 2.0%; Score 23; DB 9; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1087 TGGATCTTGGGAGATGTTTCAT 1109
|||||
Db 332 TGGATCTTGGGAGATGTTTCAT 310

Search completed: February 5, 2004, 08:12:44
Job time : 2894 secs

THIS PAGE BLANK (USPTO)